DOCUMENT RESUME

ED 355 266 TM 019 611

TITLE	Oversight Hearings on the Report of the National Council on Education Standards and Testing. Hearings before the Subcommittee on Elementary, Secondary, and Vocational Education of the Committee on Education and Labor. House of Representatives, One Hundred Second Congress, Second Session (February 4, 19, and March 18, 1992).
INSTITUTION	Congress of the U.S., Washington, DC. House Subcommittee on Elementary, Secondary, and Vocational Education.
REPORT NO	ISBN-0-16-039802-9
PUB DATE	92
NOTE	290p.; Serial No. 102-105.
AVAILABLE FROM	U.S. Government Printing Office, Superintendent of Documents, Congressional Sales Office, Washington, DC 20402.
PUB TYPE	Legal/Legislative/Regulatory Materials (090)
EDRS PRICE	MF01/PC12 Plus Postage.
DESCRIPTORS	Academic Standards; Educational Assessment; *Educational Objectives; Educational Practices; Elementary Secondary Education; *Federal Legislation; Government Role; Guidelines; Hearings; *National Competency Tests; *Student Evaluation; Testing Problems; *Test Use
IDENTIFIERS	Congress 102nd; *National Council on Educ Standards and Testing; National Education Goals Panel; *Standard Setting

ABSTRACT

The Subcommittee on Elementary, Secondary, and Vocational Education of the House Committee on Education and Labor met to consider the recommendations of the National Council on Education Standards and Testing. The report recommends the establishment of national education standards, a national system of assessments, and the establishment of a reconfigured National Education Goals Panel and a national education standards and assessment council to coordinate the development of the standards and assessments. Testifying on behalf of the Council's recommendations were M. S. Smith (Stanford University) and (L. B. Resnick, Learning Research and Development Center, University of Pittsburgh and National Council on Education Standards and Testing). Opposition to the proposal for national testing was expressed by L. Darling-Hammond (National Center for Restructuring Education, Schools and Teaching, Columbia University). W. M. Haney (Center for the Study of Testing, Evaluation, and Educational Policy, Boston College) outlined a number of shortcomings in the Council report, largely in the area of national testing. Additional statements were offered by: (1) N. V. Cantu, Mexican American Legal Defense and Educational Fund, Texas; (2) L. Rezmierski, superintendent of Northville schools, Michigan; (3) H. D. Hoover, Iowa Basic Skills Testing, Iowa; (4) M. J. Feuer, Office of Technology Assessment, accompanied by N. Carson, Office of Technology Assessment; (5) A. Shanker, American Federation of Teachers, AFL-CIO; (6) D. T. Kearns, U.S. Department of Education, Washington, D.C.; (7) R. Romer, Governor of Colorado, Co-Chair National Council on Education Standards and Testing; (8) K. Geiger, National Education Association, National Council on Education Standards and Testing; (9) B. Rosenberg, American Federation of Teachers, AFL-CIO; (10) D. M. Koretz, Rand Corporation; and (11) M. H. Kean, Association of American Publishers and CTB Macmillan/McGraw-Hill. Prepared statements by these speakers and additional prepared statements and supplemental materials are provided. (SLD)

ED355266

OVERSIGHT HEARINGS ON THE REPORT OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING

HEARINGS

BEFORE THE

SUBCOMMITTEE ON ELEMENTARY, SECONDARY, AND VOCATIONAL EDUCATION

OF THE

COMMITTEE ON EDUCATION AND LABOR HOUSE OF REPRESENTATIVES

ONE HUNDRED SECOND CONGRESS

SECOND SESSION

HEARINGS HELD IN WASHINGTON, DC, FEBRUARY 4, 19, AND MARCH 18, 1992

Serial No. 102-105

Printed for the use of the Committee on Education and Labor

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

U.S. GOVERNMENT PRINTING OFFICE

52-840 cc WASHINGTON : 1992

For sale by the U.S. Government Printing Office
Superintendent of Documents, Congressional Sales Office, Washington, DC 20402
ISBN 0-16-039802-9

TM019611

BEST COPY AVAILABLE

COMMITTEE ON EDUCATION AND LABOR

WILLIAM D. FORD, Michigan, *Chairman*

JOSEPH M. GAYDOS, Pennsylvania
WILLIAM (BILL) CLAY, Missouri
GEORGE MILLER, California
AUSTIN J. MURPHY, Pennsylvania
DALE E. KILDEE, Michigan
PAT WILLIAMS, Montana
MATTHEW G. MARTINEZ, California
MAJOR R. OWENS, New York
CHARLES A. HAYES, Illinois
CARL C. PERKINS, Kentucky
THOMAS C. SAWYER, Ohio
DONALD M. PAYNE, New Jersey
NITA M. LOWEY, New York
JOLENE UNSOELD, Washington
CRAIG A. WASHINGTON, Texas
JOSÉ E. SERRANO, New York
PATSY T. MINK, Hawaii
ROBERT E. ANDREWS, New Jersey
WILLIAM J. JEFFERSON, Louisiana
JOHN F. REED, Rhode Island
TIM ROEMER, Indiana
JOHN W. OLVER, Massachusetts
ED PASTOR, Arizona
RON DE LUGO, Virgin Islands
JAIME B. FUSTER, Puerto Rico

WILLIAM F. GOODLING, Pennsylvania
E. THOMAS COLEMAN, Missouri
THOMAS E. PETRI, Wisconsin
MARGE ROUKEMA, New Jersey
STEVE GUNDERSON, Wisconsin
RICHARD K. ARMEY, Texas
HARRIS W. FAWELL, Illinois
PAUL B. HENRY, Michigan
CASS BALLENGER, North Carolina
SUSAN MOLINARI, New York
BILL BARRETT, Nebraska
JOHN A. BOEHNER, Ohio
SCOTT L. KLUG, Wisconsin
MICKEY EDWARDS, Oklahoma
RANDY "DUKE" CUNNINGHAM, California

PATRICIA F. RISSLER, *Staff Director*
JAY EAGEN, *Minority Staff Director*

SUBCOMMITTEE ON ELEMENTARY, SECONDARY, AND VOCATIONAL EDUCATION

DALE E. KILDEE, Michigan, *Chairman*

GEORGE MILLER, California
PAT WILLIAMS, Montana
MATTHEW G. MARTINEZ, California
CARL C. PERKINS, Kentucky
CHARLES A. HAYES, Illinois
THOMAS C. SAWYER, Ohio
MAJOR R. OWENS, New York
NITA M. LOWEY, New York
JOLENE UNSOELD, Washington
WILLIAM J. JEFFERSON, Louisiana
JOHN F. REED, Rhode Island
TIM ROEMER, Indiana
CRAIG A. WASHINGTON, Texas
PATSY T. MINK, Hawaii
JOHN W. OLVER, Massachusetts
ED PASTOR, Arizona
JAIME B. FUSTER, Puerto Rico

WILLIAM F. GOODLING, Pennsylvania
SCOTT L. KLUG, Wisconsin
THOMAS E. PETRI, Wisconsin
MARGE ROUKEMA, New Jersey
STEVE GUNDERSON, Wisconsin
PAUL B. HENRY, Michigan
SUSAN MOLINARI, New York
JOHN A. BOEHNER, Ohio
MICKEY EDWARDS, Oklahoma
RANDY "DUKE" CUNNINGHAM, California

CONTENTS

Page

Hearings held in Washington, DC:
- February 4, 1992 1
- February 19, 1992 81
- March 18, 1992 181

Statement of:
- Cantu, Norma V., Director for Education Programs, Mexican American Legal Defense and Educational Fund, San Antonio, TX; Dr. Leonard Rezmierski, Superintendent of Northville Schools, Northville, MI, the Council for Exceptional Children; and H. D. Hoover, Director, Iowa Basic Skills Testing, Iowa City, IA 217
- Feuer, Michael J., Ph.D., Senior Analyst and Project Director, Office of Technology Assessment; accompanied by Nancy Carson, Office of Technology Assessment 82
- Kearns, Hon. David T., Deputy Secretary of Education, U.S. Department of Education, Washington, DC 182
- Romer, Hon. Roy, Governor, State of Colorado, Co-Chair, National Council on Education Standards and Testing; Keith Geiger, President, National Education Association and Member, National Council on Education Standards and Testing; Bella Rosenberg, Assistant to the President, American Federation of Teachers, AFL-CIO; Daniel M. Koretz, Ph.D., Senior Social Scientist, Rand Corp.; and Michael H. Kean, Ph.D., Chair, Test Committee, Association of American Publishers and Vice President, Public and Governmental Affairs, CTB Macmillan/McGraw-Hill 109
- Smith, Marshall S., Dean, School of Education, Stanford University, Member, National Council on Education Standards and Testing; Dr. Lauren B. Resnick, Director, Learning Research and Development Center, University of Pittsburgh, Member, National Council on Education Standards and Testing; Dr. Linda Darling-Hammond, Co-director, National Center for Restructuring Education, Schools and Teaching, Columbia University; and Dr. Walter M. Haney, Senior Research Associate, Center for the Study of Testing, Evaluation, and Educational Policy, Boston College 3

Prepared statements, letters, supplemental materials, et cetera:
- Baker, Eva L., Professor, University of California, Los Angeles and Robert L. Linn, Professor, University of Colorado, prepared statement of 275
- Cantu, Norma V., Director for Education Programs, Mexican American Legal Defense and Educational Fund, San Antonio, TX, prepared statement of 221
- Darling-Hammond, Dr. Linda, Co-director, National Center for Restructuring Education, Schools and Teaching, Columbia University, prepared statement of 31
- Feuer, Michael J., Ph.D., Senior Analyst and Project Director, Office of Technology Assessment, prepared statement of 87
- Geiger, Keith, President, National Education Association and Member, National Council on Education Standards and Testing, prepared statement of 118
- Goodling, Hon. William F., a Representative in Congress from the State of Pennsylvania, prepared statement of 82
- Haney, Dr. Walter M., Senior Research Associate, Center for the Study of Testing, Evaluation, and Educational Policy, Boston College, prepared statement of 41

Page

Prepared statements, letters, supplemental materials, et cetera—Continued
Hoover, H.D., Director, Iowa Basic Skills Testing, Iowa City, IA, prepared statement of 241
Kean, Michael H., Ph.D., Chair, Test Committee, Association of American Publishers and Vice President, Public and Governmental Affairs, CTB Macmillan/McGraw-Hill, prepared statement of 155
Kearns, Hon. David T., Deputy Secretary of Education, U.S. Department of Education, Washington, DC, prepared statement of 185
Koretz, Daniel M., Ph.D., Senior Social Scientist, Rand Corp., prepared statement of 135
Lieberman, Dr. Ann, President, AERA, Teachers College, Columbia University, prepared statement of 263
Martinez, Hon. Matthew G., a Representative in Congress from the State of California, prepared statement of 63
National Alliance of Business, prepared statement of 265
National Association of State Boards of Education, Alexandria, VA, prepared statement of 282
Resnick, Dr. Lauren B., Director, Learning Research and Development Center, University of Pittsburgh, Member, National Council on Education Standards and Testing, prepared statement of 21
Rezmierski, Dr. Leonard, Superintendent of Northville Schools, Northville, MI, the Council for Exceptional Children, prepared statement of 231
Romer, Hon. Roy, Governor, State of Colorado, Co-Chair, National Council on Education Standards and Testing, prepared statement of 112
Sawyer, Hon. Thomas C., a Representative in Congress from the State of Ohio, prepared statement of 58
Shanker, Albert, President, American Federation of Teachers, AFL-CIO, prepared statement of 125
Smith, Marshall S., Dean, School of Education, Stanford University, Member, National Council on Education Standards and Testing, prepared statement of 9
Stewart, Donald M., President, The College Board, Washington, DC, prepared statement of 75

OVERSIGHT HEARING ON THE REPORT OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING

TUESDAY, FEBRUARY 4, 1992

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON ELEMENTARY, SECONDARY,
AND VOCATIONAL EDUCATION,
COMMITTEE ON EDUCATION AND LABOR,
Washington, DC.

The subcommittee met, pursuant to call, at 9:30 a.m., Room 2261, Rayburn House Office Building, Hon. Dale Kildee, Chairman, presiding.

Members present: Representatives Kildee, Miller, Williams, Martinez, Hayes, Sawyer, Owens, Lowey, Reed, Roemer, Pastor, Goodling, Petri, Roukema, Gunderson, and Molinari.

Staff present: Susan Wilhelm, staff director; Jeff McFarland, legislative counsel; Damian Thorman, legislative associate; Jack Jennings, general counsel (education); Andrew Hartman, minority education coordinator; and Lynn Selmser, minority professional staff member.

Chairman KILDEE. The Subcommittee on Elementary, Secondary, and Vocational Education convenes today for the first of several hearings on the recommendations of the National Council on Education Standards and Testing.

Before I proceed, I would like to call to the attention of those who are assembled here today that one of my mentors and the former Chairman of the full committee—a person who I can truly say, as God as my witness, that I'm a better person because of him—is present here this morning, the honorable Gus Hawkins from California.

As members of the committee are aware, the Council was established by legislation originating in this committee to advise the Congress, the Secretary of Education, and the National Goals Panel, on the desirability and feasibility of national education standards and national assessments. The Council released its report January 24th, and copies have been distributed to all members.

The report recommends the establishment of national education standards, a national system of assessments, and the establishment of a reconfigured national education goals panel and a national education standards and assessment council, to coordinate the development of both the standards and assessments.

The topics addressed in this report and the recommendations it makes must be considered carefully and deliberately. Our goal must be to determine the most effective way to improve our education system and to ensure the academic success of all children.

Testifying on behalf of the Council this morning are Dr. Marshall "Mike" Smith, Dean of the College of Education at Stanford University—and Mike and I got to know each other very well during those meetings—and Dr. Lauren B. Resnick, Director of the Learning Research and Development Center, the University of Pittsburgh. Lauren and I also got to know each other and shared many good ideas during that time. Mike and Lauren are both active members of the Council.

They are joined by Dr. Linda Darling-Hammond, Co-Director, National Center for Restructuring Education, Schools and Teaching, Columbia University, and Dr. Walter M. Haney, Senior Research Associate, Center for the Study of Testing, Evaluation, and Educational Policy, Boston College.

Before we begin, I would like to recognize my good friend, a good educator, the ranking Republican on both this subcommittee and the full committee, Mr. Goodling.

Mr. GOODLING. Who as of midnight last night was voted back a district which I had lost just a few days before.

[Laughter.]

Chairman KILDEE. I think you're doing better than I am right now.

Mr. GOODLING. All I want to say is that the only person I like better than our "Chairman Gus" is Elsie. I, too, welcome you here.

I hope some of our other colleagues get here because this is the church talking to the choir, or the choir talking to the church, something of that nature. I'm glad we got one to join us, and hopefully some others will come, although maybe the White House would like it to be the way it is, because they couldn't imagine that we couldn't push it through here as rapidly as it was pushed through the Senate. I said, well, they had no idea what it was they were voting on, but I'll guarantee you, by the time we get to it, the House Members will know and it may be more difficult.

Chairman KILDEE. Thank you.

Mr. Roemer.

Mr. ROEMER. Thank you, Mr. Chairman. I, too, would like to welcome Chairman Hawkins to this hearing this morning and salute both you, Mr. Chairman, as well as the ranking member, for the quick start here to have hearings on such an important topic. I look forward to the testimony from our distinguished panelists this morning.

I would also like to say that 2 weeks ago, Mr. Chairman, I testified before the Joint Economic Committee and talked about education as a component of a "Marshall Plan" to rebuild and restructure America. Education is such an important pillar of this Marshall Plan to rebuild. I am hopeful to hear some innovative, creative, new approaches to education this morning for our country. We need to tie new Federal moneys to innovation and genius in this system, and I look forward to hearing some of those ideas this morning.

Thank you, Mr. Chairman.

Mr. GOODLING. I would like to say that I watched you as you were testifying, and I particularly liked your line when they said something about the money, and you said "well, you've been here a lot longer than I have; it's up to you to find that money."

Mr. ROEMER. Thank you. Every now and then I do come up with a good line.

Chairman KILDEE. Mike, you may begin your testimony this morning.

STATEMENTS OF MARSHALL S. SMITH, DEAN, SCHOOL OF EDUCATION, STANFORD UNIVERSITY, AND MEMBER, NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING; DR. LAUREN B. RESNICK, DIRECTOR, LEARNING RESEARCH AND DEVELOPMENT CENTER, UNIVERSITY OF PITTSBURGH, AND MEMBER, NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING; DR. LINDA DARLING-HAMMOND, CO-DIRECTOR, NATIONAL CENTER FOR RESTRUCTURING EDUCATION, SCHOOLS AND TEACHING, COLUMBIA UNIVERSITY; AND DR. WALTER M. HANEY, SENIOR RESEARCH ASSOCIATE, CENTER FOR THE STUDY OF TESTING, EVALUATION, AND EDUCATIONAL POLICY, BOSTON COLLEGE

Mr. SMITH. Mr. Chairman, Congressman Goodling, Congressman Roemer, I am honored to be here today to discuss with you the very important issues of national standards, national testing, and the quality of education in the United States for all of our students.

I am going to suppress the urge to have a pun about a "Marshall Plan." I am speaking here as a member of the National Council on Education Standards and Testing, and as the chair of the Task Force on Standards of the Council.

I will speak about three aspects of the Council's work, and would be happy to answer questions about other aspects.

The three aspects are the overall recommendations of the Council, the specific recommendations about the desirability and feasibility of national standards. This includes ways in which issues of equal education opportunity are addressed in the report. I'm going to leave it to Dr. Resnick to focus on the issues of assessment.

The third issue that I will discuss is the nature of a mechanism for helping to implement the recommendations of the Council.

The Council was a diverse body. There were 32 members. The members included a broad spectrum of the education world, including a "Teacher of the Year," two superintendents, two principals, as well as representatives of State and the Federal Government from both parties. The Council met seven times, heard over 30 presentations, solicited and received comments from a large number of organizations, and considered reports generated by eight task forces, each chaired by a member of the Council.

Decisions were hard to come by in the Council, I think would be a fair thing to say. The Council was vigorous in its analysis of the charge given to it by the Congress. There were many disagreements along the way, and a number of compromises are reflected in the report. The issues set out in the legislation concerning desir-

ability and feasibility of national education standards and testing were debated and redebated during the life of the Council.

Yet, even though there were initial disagreements, the Council concluded that, under the right conditions, voluntary national standards are desirable and feasible; and second, that a single system, not a single test, of voluntary national examinations based on the national standards is desirable and feasible.

Finally, the Council recommended that a national coordinating structure be put in place to advance standard setting and assessment development as part of a comprehensive educational reform.

These may appear on the surface to be innocuous recommendations. They don't ask for big money to go into education; they don't suggest ways of changing specific activities in a classroom. But they are not innocuous recommendations. If implemented, they will represent a major departure in the United States' way of doing business in education. For the first time, this Nation will develop a set of education standards and performance expectations for all of its children. In a dramatic and important way, we would be explicitly recognizing that the quality and equality of our country's system of education for all students is a national as well as a State and local responsibility, and we would be taking a significant step toward improving that education.

Now, how did we conclude that voluntary national standards are desirable and feasible? We had to go through a number of steps. The first step was to define what national standards are—and I'll read you the definitions that we ended up with. A second step was to consider under what conditions national standards might be desirable and feasible, and then the third step was to answer the three questions that the Congress posed to us: about whether or not national standards would influence achievement and teaching in the classroom; about the influence of national standards on educational equity; and about the role of national standards, national standards, in as diverse a Nation as we presently have. That's what I will cover in the next 2 or 3 minutes.

What are education standards and under what conditions are they desirable and feasible? The Council recommends standards for students and standards for schools and school systems. Standards for students are more common, more understood, I believe. They have two parts:

Content standards for students describe the knowledge, skills and common understandings that students are expected to learn and, thus, that schools are expected to teach.

Student performance standards, the second set of standards for students, define levels of competence in the challenging subject matter set out in the content standards.

Now, we can specify the content standards, and we can set ourselves expectations about performance standards. But it all doesn't mean anything unless we begin to think about what it takes to deliver the education, to deliver the teaching, to provide the opportunity to learn to students so that they can accomplish, so they can learn those content standards, at very high levels specified by the performance standards.

So, to ensure that students do not bear the sole burden of attaining the standards, and to encourage assurances that all students

will have the opportunity to learn the content standards to a high level of performance, the Council recommends two forms of standards for the education system.

School delivery standards would be established by the States for the purpose of assessing whether a school delivers to all students the opportunity to learn the content standards.

System delivery standards would be established to provide evidence about the success of schools, local school systems, States and the Nation, in bringing all students to high performance standards.

These are very important steps that the Council took. This is an example of a set of issues that were debated and redebated over time, over a 6 month's period of time. They're terribly important and they really go to the heart of thinking about major educational reform in the United States.

Now, in its discussions, the Council established a set of necessary conditions for finding that national educational standards are desirable and feasible. There are five such conditions:

First, the standards should generate high expectations, not expectations of minimal competency. Second, they should provide focus and direction, not a national curriculum. Third, they should be national, not Federal. They should not be controlled by the Federal Government. They should be controlled by nationally represented bodies. Fourth, they should be voluntary, not mandatory. And fifth, they should be dynamic, not static. They should change as our knowledge and understandings change over time.

These definitions and conditions set the stage for the Council's consideration of the three questions I posed earlier. The key questions about desirability posed by the Congress in legislation creating the Council.

The first question was, will national standards have a positive influence on student achievement and the quality of teachers and schools? This is not a simple question, either. None of these questions were simple.

If there were voluntary, challenging national standards which had been developed through a broad national consensus process, the Council found that one of the primary conditions for improving the Nation's performance would have been met; that is, the Nation would have changed its expectations from minimal standards to higher standards. But even if that step were taken, the hard work will only have begun.

Standards alone cannot change student achievement, school quality, or teacher performance. Student achievement and teacher performance will not be greatly influenced by content and performance and delivery standards unless the standards are part of a coherent and systemic approach to improving instruction in the schools. This approach would have a number of components:

First, it would have challenging national standards which could set expectations for all schools and grades in key content areas. Second, it would have a voluntary national examination system based on the content standards that could reinforce and assess attainment of the standards. Third, the challenging content and performance standards could form the basis for systemic reform of State and local policies regarding the adoption of instructional materials and teacher professional development and licensing. Fourth,

the school system delivery standards would establish fundamental criteria by which to assess the quality and capacity of all of our schools to deliver to all of our children the opportunity to learn at a high level of performance the material in the content standards. There would be no excuse for settling for a less rigorous curriculum for some children. These arguments were persuasive to the Council.

The second question, will national standards improve or harm educational equity for students? The Council was unanimous in its view that a major part of the justification for national standards must rest on their promise for improving the quality of the educational experiences for the most needy in our society.

This was also a complex issue, discussed in great detail throughout the life of the Council.

Three arguments stand out: first, common challenging national standards and assessments for all can be a powerful catalyst for implementing the systemic change necessary to bring all students to high performance levels. These standards would apply to all children. In my community, they would apply to East Palo Alto as well as Palo Alto. Other communities have other examples of the same sort. Right now we do not have common standards applying at Palo Alto and to East Palo Alto. The standards in Palo Alto are far higher, reflecting a far more advantaged population.

Second, school delivery standards would, for the first time, establish fundamental criteria for assessing whether schools can provide to all children the opportunity to learn the challenging content standards to a high level of performance.

Third, the proposed examination system could be used as a mechanism for helping to ensure equal opportunity to learn for all children. The Council was clear in its recommendations that the examination system proposed by the Council could only be used for the purpose of helping make decisions about students and schools if the examinations met the standards of high reliability and validity.

The validity standard is critical. The proposed examinations are intended to assess student performance on the content standards. For such an examination to be valid, the student must have had the opportunity to learn the material on the examination. This, in turn, implies that the school has the capacity to deliver to all students the opportunity to learn the material to a high performance level. Otherwise, the examination would not be valid and could not be legitimately used for decision-making purposes.

During the discussions of equal opportunity in the Council, I often recalled the real-life experience of the primarily low income, Mexican-American students depicted in the movie "Stand and Deliver." These students were exposed, many for the first time in their lives, to a common and, indeed, national set of very challenging content and performance standards. Their level of academic performance was then assessed by the national advanced placement examinations in calculus.

The students performed extraordinarily well. They responded to the clear challenge of the content usually reserved for students from much more advantaged backgrounds. But they were also well prepared by a master teacher, with deep knowledge of the content and great pedagogical skill. All of these ingredients are necessary

for our students, rich or poor, to succeed in raising their achievement performance—clear content standards, clear performance standards, and a school and system and teacher prepared to give them the opportunity to learn.

The third question posed by the Congress had to do with the issue of whether national standards are appropriate for a Nation with very diverse populations, where the curriculum is traditionally controlled at the State and local level.

To a substantial degree, these issues are met by the two conditions already stated. First, the standards are voluntary, and in being voluntary, local and State school systems can either choose and select them or not select them. Second, they serve only as a common core around which States, localities and schools maintain their flexibility and responsibility of tailoring their specific curricula to meet the particular needs of their students.

A second concern raised by these issues is the effect of the wide variations that exist in State and local fiscal and human resources. This is an important concern, for the variation in resources may well contribute to variation in performance and delivery standards among and within States.

The simple fact, however, is that variations in resources should not be used to justify and excuse wide variation in the quality of content presented or the levels of student performance, as now occurs. Well defined and challenging national standards could serve both to point out problems and to establish clear targets for all States and localities.

Based on these considerations, the Council concluded that the development of national education standards would be desirable. If properly designed and embedded within a larger system of education reform, the standards could lead to higher quality teaching and learning, a more equitable educational environment, and would enhance the sense of a common identity and community in the Nation.

The final question I want to address is should there be a national coordinating structure to advance standard setting and assessment development.

The Council recommended that a coordinating structure at the national level be created to help organize the development and implementation of voluntary national standards and a system of assessments. Such a structure would serve three critical purposes:

First, it would help maintain the momentum developed by national public interest and through the work of the Council and many professional groups, such as the NCTM.

Second, it would help ensure that mass confusion at the national and Federal level is not created as different groups and agencies respond to the national interest in education standards and assessment. We need a coordinating body just in order to keep track of what's going on and to impose some quality control.

Third, it would help to ensure that there is public involvement and accountability as the process of standard setting and examination development proceeds.

Conceived to be both independent of the Federal Government and accountable to the public, such a structure could serve to help

mobilize the entire country and move toward challenging standards and a high quality education experience for all children.

Let me close by quoting a brief statement from the Council report, a statement that occurs in its concluding section on ensuring equity in education:

"Providing genuine opportunity for all students to achieve high standards is a moral imperative. The standards that the Council proposes would apply to the entire education system. All students must have the opportunity to achieve them and to be assessed fairly on their attainment. To bring this about, equitable educational opportunities must be provided.

"The Council recognizes the concerns of those who are fearful of the unintended consequences of its proposals. Yet, high standards and knowledge gained from appropriate assessments could serve as rallying points to secure the school and community efforts to reach them. High-quality standards and assessments should mobilize educators and the public to reform schools, engage families and communities, create incentives for high performance, and provide genuine opportunity for all students."

This concludes my testimony. I would be happy to respond to questions.

[The prepared statement of Marshall S. Smith follows:]

OVERSIGHT HEARING ON THE REPORT OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING

COMMITTEE ON EDUCATION AND LABOR

U.S. HOUSE OF REPRESENTATIVES

SUBCOMMITTEE ON ELEMENTARY, SECONDARY AND VOCATIONAL EDUCATION

Marshall S. Smith, Dean
School of Education
Stanford University
Stanford, CA 94305-3096

February 4, 1992

Testimony before the House Subcommittee on Elementary and Secondary Education on the Report of The National Council on Education Standards and Testing.

Marshall S. Smith
Dean, School of Education
Stanford University

Chairman Kildee, Congressman Goodling, other Members of Congress:

I am honored to be here today to discuss with you the very important issues of national standards, national testing, and the quality of education in the United States for all of our students.

In the legislation creating the National Council on Education Standards and Testing, Congress asked the Council to consider the "desirability" and the "feasibility" of national standards and tests. In addition, the legislation requested the Council to make recommendations about "long-term policies, structures, and mechanisms for setting voluntary education standards and an appropriate system of tests."

I was a member of the National Council on Education Standards and Testing and the Chair of the Council's Task Force on Standards. I will speak about three aspects of the Council's work and would be happy to answer questions about other aspects. The three aspects are:

1. The overall recommendations of the Council.

2. The specific recommendations about the desirability and feasibility of national standards. This includes ways in which issues of equal educational opportunity are addressed in the report.

3. The nature of a "mechanism" for helping to implement the recommendations of the Council.

Council Recommendations:

There were 32 members of the Council. The members included a broad spectrum of the education world including a teacher of the year, two superintendents, and two principals as well as representatives of state and the federal governments from both parties. The Council met seven times, heard over thirty presentations, solicited and received comments

from a large variety of organizations, and considered reports generated by eight task forces, each chaired by a member of the Council.

The Council was unanimous in its recommendations. All members of the Council concur with the report, Raising Standards for American Education. Unanimity was not arrived at easily -- there were many disagreements along the way and a number of compromises are reflected in the report. The issues set out in the legislation concerning the feasibility and desirability of national education standards and testing were debated and re-debated during the life of the Council.

Yet even though there were initial disagreements, all members of the Council concluded that, under the right conditions:

1. **voluntary national standards are desirable and feasible and;**

2. **a system (not a single test) of voluntary national examinations based on the national standards is desirable and feasible.**

Finally, the Council recommended that a national "coordinating structure be put in place to advance standard-setting and assessment development ... as part of a comprehensive educational reform."

These may appear to some to be innocuous recommendations. They are not. If implemented they will represent a major departure in the United State's way of doing business in education. For the first time, this **nation** would develop a set of education standards and performance expectations for all of our children. In a dramatic and important way, we would be explicitly recognizing that the quality and equality of our country's system of education for all students is a national as well as a state and local responsibility, and we would be taking a significant step toward improving that education.

Challenging, Voluntary, National Education Standards:

What are education standards and under what conditions are they desirable and feasible? The Council recommends standards for students and standards for schools and school systems. Standards for students have two parts:

1. **Content Standards** describe the knowledge, skills and common understandings that students are expected to learn and, thus, that schools are expected to teach. In some states, content standards are termed "curriculum frameworks" because they provide the broad focus and direction around which local curriculum is developed.

2. **Student Performance Standards** define levels of competence in the challenging subject matter set out in the content standards.

To ensure that students do not bear the sole burden of attaining the standards and to encourage assurances that all students will have the opportunity to learn the content standards to a high level of performance, the Council recommends two forms of standards for the education system.

3. **School Delivery Standards** would be established by the states for the purpose of assessing whether a school delivers to all students the opportunity to learn the content standards.

4. **System Delivery Standards** would be established to provide evidence about the success of schools, local school systems, states and the nation in bringing all students to high performance standards.

In its discussions the Council established a set of necessary conditions for finding that national education standards are desirable and feasible. National standards should meet the following conditions:

1. **High Expectations -- not expectations of minimal competency.** The content, performance, and delivery standards should represent **challenging** levels for all students and schools to attain.

2. **Focus and Direction -- not a national curriculum.** The content standards should provide a common core of knowledge, skills and understandings -- a core which would be augmented and enhanced by state and local flexibility.

3. **National -- not federal.** The standards should be arrived at through a broad national effort, not created and/or mandated by the federal government. The effort should involve the widest possible participation from individuals and groups at the national, state, and local levels.

4. **Voluntary -- not mandatory.** States and local agencies would not have to adopt the national standards. The Council expected there would be widespread adoption if the standards were of the highest quality.

5. **Dynamic -- not static.** While care should be taken to develop quality standards from the outset, the process must be on-going, entailing continuous improvement over time.

These definitions and conditions set the stage for the Council's consideration of three key questions about desirability posed by the Congress in the legislation creating the Council.

Will National Standards have a positive influence on student achievement and the quality of teachers and schools?

The Council found that at the present time our nation's de facto content, performance and delivery standards are mediocre at best. Moreover, it is difficult to imagine improving our level of student achievement or the quality of our schools by very much if we continue to support and implement our current minimal standards curricula.

If there were voluntary, challenging national standards which had been developed through a broad national consensus process, one of the primary conditions for improving the nation's performance would have been met: the nation would have changed its expectations. But, the hard work will have only begun.

Standards alone cannot change student achievement, school quality, or teacher performance. Student achievement and teacher performance will not be greatly influenced by content and performance and delivery standards unless the standards are part of a coherent and systemic approach to improving instruction in the schools. This approach would have a number of components:

- **Challenging national standards could set expectations for all schools and grades in key content areas, signalling the type of substantive changes we need system-wide in our schools and classrooms.**

- **A voluntary national examination system based on the content standards could reinforce and assesses attainment of the standards.**

- **The challenging content and performance standards could form the basis for systemic reform of state and local policies regarding the adoption of instructional materials and teacher professional development and licensing. We would then have several interconnected policy efforts directed toward improved teaching and learning around ambitious, not minimal, academic goals.**

- **The school and system delivery standards would establish fundamental criteria by which to assess the quality and capacity of our schools to deliver to all children the opportunity to learn to a high level of performance the material in the content standards. There would be no excuse for settling for a less rigorous curriculum for some children.**

These arguments were persuasive to the Council. We would not expect immediate improvement in student achievement and school quality. Indeed, as these points indicate the path will be very demanding. But, we believe it begins with a challenging, coherent, and common vision of what our goals and standards are -- we need to know where we want to go before we can set out on the path.

Will National Standards improve or harm Educational Equity for Students?

The Council was unanimous in its view that a major part of the justification for

national standards must rest on their promise for improving the quality of the educational experiences for the most needy in our society.

The Council addressed this issue throughout the report, not just in its consideration of national standards. Three arguments stand out:

- **Common, challenging national standards and assessments for all can be a powerful catalyst for implementing the systemic change necessary to bring all students to high performance levels.** By emphasizing their applicability to all students, national standards and assessments could help assure that adequate resources are available and appropriately targeted to helping all students attain the standards.

- **School Delivery standards would, for the first time, establish fundamental criteria for assessing whether a schools can provide to all children the opportunity to learn the challenging content standards to a high level of performance.** These criteria would not be geared to minimal de facto expectations and standards -- they would be developed in the context of challenging content and performance standards for all children.

- **The proposed examination system can be used as a mechanism for helping to insure equal opportunity to learn for all children.** The Council was clear in its recommendations that the examination system proposed by the Council could only be used for the purpose of helping to make decisions about students and schools if the examinations met the standards of high reliability and validity. The validity standard is critical. The proposed examinations are intended to assess student performance on the content standards. For such an examination to be valid, the student must have had the opportunity to learn the material on the examination. This in turn implies that the school has the capacity to deliver to all students the opportunity to learn the material to a high performance level -- otherwise the examination would not be valid and could not legitimately be used for decision making purposes.

During the discussions of equal opportunity in the Council, I often recalled the real-life experience of the primarily low income, Mexican American students depicted in the movie "Stand and Deliver". These students were exposed, many for the first time in their lives, to a common (indeed, national) set of very challenging content and performance standards. Their level of academic performance was then assessed by the national Advanced Placement Examinations in calculus. The students performed extraordinarily well -- they responded to the clear challenge of the content usually reserved for students from much more advantaged backgrounds. But they were also well prepared by a master teacher with deep knowledge of the content and great pedagogical skill. All of these ingredients are necessary for our students, rich or poor, to succeed in raising their achievement performance -- clear content standards, clear performance standards, and a school and system prepared to give to them the opportunity to learn. The delivery standards recommended by the Council would help to ensure that schools and school systems provide that opportunity.

Are National Standards appropriate for a nation with very diverse populations where the curriculum is traditionally controlled at the state and local level?

To a substantial degree these issues are addressed by two conditions: first, that the standards are voluntary, and second, that they serve only as a common core around which states, localities and schools maintain the flexibility and responsibility of tailoring their specific curricula to meet the particular needs of their students. Moreover, the Council believed that a national consensus around common standards could enhance the sense of national identity and community we need as our Nation becomes increasingly diverse. National polls reflect overwhelming support among all groups for the concept of national education standards.

A second concern raised by these issues is the effect of the wide variations that exist in state and local fiscal and human resources. This is an important concern, for the variation in state and local resources may well contribute to variation in performance and delivery standards among and within states.

The simple fact, however, is that variations in resources should not be used to justify and excuse wide variation in the quality of content presented or the levels of student performance, as now occurs. Well defined and challenging national standards (content, performance, and delivery) could serve both to point out problems and to establish clear targets for all states and localities to strive for.

Summary:

Based on these considerations, the Council concluded that the development of National Education Standards would be desirable. If properly designed and embedded within a larger system of education reform the standards would lead to higher quality teaching and learning, a more equitable educational environment and would enhance the sense of a common identity and community in the nation.

The Council also considered the feasibility of establishing challenging, national standards. It concluded that there was considerable evidence that very high quality content standards could be developed. Indeed, the nation already has such standards in mathematics and a number of states are moving to develop content standards in other academic areas.

The U. S. experience with student performance standards is less extensive but the work of the National Assessment Board and the College Board on the Advanced Placement Examinations gives the nation a good beginning. The design of appropriate school and system delivery standards in the context of common content standards is even newer ground for the U. S. though again the mathematics community through the NCTM guidelines for instruction is leading the way.

Overall, then, the Council concluded that national education standards were both desirable and feasible.

Should there be a national coordinating structure to advance standards-setting and assessment development?

The Council recommended that a coordinating structure at the national level be created to help organize the development and implementation of voluntary national standards and a system of assessments. The Council believed that the implementation of national standards and assessments should not take place in isolation "but should be part of comprehensive education reform. Indeed, the intellectual and political activities of setting high national standards and developing assessments are likely to provide added momentum for higher expectations and educational renewal."

Such a structure would serve three critical purposes. First, it will help maintain the momentum developed by national public interest and through the work of Council and many professional education groups such as the NCTM. Second, it will help insure that mass confusion at the national and federal level is not created as different groups and agencies respond to the national interest in education standards and assessment. Without such a structure to help coordinate response, we could see the generation of uncoordinated piecemeal procurements of standards and examinations in different areas based on different standards of quality. Third, it will help to insure that there is public involvement and accountability as the process of standard setting and examination development proceeds.

Conceived to be both independent of the federal government and accountable to the public, such a structure could serve to help mobilize the entire country to move toward challenging standards and a high quality education experience for all children.

The Council report summarizes the position of the Council in its concluding section on "Ensuring Equity":

> Providing genuine opportunity for all students to achieve high standards is a moral imperative. The Standards that the Council proposes would apply to the entire education system. All students must have the opportunity to achieve them and to be assessed fairly on their attainment. To bring this about, equitable educational opportunities must be provided.
>
> The Council recognizes the concerns of those who are fearful of the unintended consequences of its proposals. Yet high standards and knowledge grained from appropriate assessments could serve as rallying points to secure the school and community efforts to reach them. High-quality standards and assessments should mobilize educators and the public reform schools, engage families and communities, create incentives for high performance, and provide genuine opportunity for all students."

7

Chairman KILDEE. Thank you very much.

Dr. Resnick.

Dr. RESNICK. I am very pleased to be here. I come from the University of Pittsburgh, where I am Director of the Learning Research and Development Center. Like Mike, I was a member of the National Council on Education Standards and Testing and am a supporter of its recommendations.

This morning I would like to say why I believe that those recommendations are ones that Congress should act on, especially when joined with the provisions of the Neighborhood Schools Improvement Act, H.R. 3320, and perhaps only when joined with provisions like those.

I have three reasons. The first and central one has to do with equity. American children today do not have an equal chance to learn in school. There are dramatic differences in resources available in rich and poor school districts, and these are growing. But the most profound inequities are often hidden; they persist even when children from different backgrounds attend the same schools. Even in such schools, expectations are not the same for all, and the opportunities for learning are not equivalent. Almost everywhere, we expect less of children from poor and minority families; we therefore ask less of them and, above all, we offer them less.

The examples are legion. There are some in my written document that have to do with the difference in de facto curricula. The weak learners who are practicing computations that they're not expected to understand, the good learners, so called, solving interesting problems, all of this in the first grade, in the same classroom, and carried forward a few years, what were initially small manageable differences have become a self-fulfilling prophecy of inability to learn on the part of some.

For poor and minority parents, these differences in expectation and, therefore, in the kind of instruction that children are offered, are nearly impossible to confront in today's system. It is not unusual to hear stories from such parents like the following, which is one that I heard recently:

A child who was earning A's and B's in a predominantly minority school moved to a more mixed school in sixth grade. There he was judged unable to read, he was given failing grades, and eventually said to be in need of remediation—to the astonishment and anger of his parents. What is happening is our schools are operating on a double standard. The people most aware of that are poor and minority families, parents especially, who have begun to get really involved in their children's education. What they recognize is that clear and public standards, the same ones for everyone, are the only way out of this situation.

Last fall, the New Standards Project, of which I'm co-director with Mort Tucker of the National Center on Education and the Economy, ran a series of focus groups in a number of locations around the country. What we were seeking was information about the opinions of different groups of Americans, people from all social and economic groups, on the desirability of national education standards and assessments. The results showed support for national standards from all groups, but the strongest and most outspoken support came from low income and minority parents who

wanted their children to be judged by the same criteria as other children, and who saw in a program of standards and assessments to go with them a chance to improve the educational prospects for their children.

So that is the primary reason that I think we need a national program of standards.

Now, the second reason, also expressed in the Council's report, is that we need to increase the education system's capacity to deliver on these standards. We are talking about something really quite radically new. We're talking not about a small rise in our current plan for what children ought to learn but a redefinition toward what some of us call a thinking curriculum for everybody.

Our schools aren't organized for that. The curriculum in place is not designed to bring students up to the kind of new standards that we need. Teachers and administrators themselves were educated in a different system. So we're all imprisoned in our past, and only a massive boot-strapping operation is going to take us where we now need to go and want to go. We're going to have to mobilize all resources and coherent programs aimed at meeting ambitious new goals.

Bootstrapping ourselves into that kind of educational future will require operating on all fronts at once—standards, assessment, curriculum and teaching materials, professional development—that may be the absolute core—preschool education, and adequate health and social services for children in school. The Standards Council recognized that, as does H.R. 3320.

Now, in addressing the desirability of assessments, the Council was clear about why a new system of assessments is desirable. It is not to monitor education achievement and measure it and show yet again that we're doing badly. We have a very fine system for doing that called the National Assessment of Educational Progress. What we need now is a tool to raise the achievement levels of children, all children, and to do that we need new forms of assessment of the following kinds:

These new assessments need to exemplify targets of instruction, making it clear to everybody—students, teachers, parents—just what kinds of capabilities students should be learning. Those assessments need to be as rich and complex as the world for which students are preparing. They need to be linked to local decisions about curriculum so that they can legitimately organize teaching and learning. They need to engage the loyalty and understanding of teachers and educators throughout the system, and they must be reliable, fair and valid for their stated purposes.

That strategy of using assessments as part of a bootstrapping operation, a lever for all the rest that needs to be done, can only work if teachers and educators are committed to assessments as educational targets. That means that a single national test cannot work. Imposed from far away, without the engagement of educators, there is no way for these assessments to penetrate educational practice.

That is why the Council was explicit in its call for a varied system of State and local assessments linked to national standards. Recent press coverage has misrepresented the Council's proposal. It has cast it as a central curriculum versus local controls. The Coun-

cil didn't call for a national test. It was quite explicit in rejecting that idea. Such a test would not serve the ends the Council had in mind. It would lock inequities into place and it wouldn't increase the education system's capacity.

So what we are looking for—here's my third point, and final one—is a high performance education system. A system of national standards, with locally developed assessments tied to them, is the key to that, key to the possibility for opening up real opportunity for all American children. That kind of system can set achievement targets that aren't just higher, but different in kind, able to produce the literate, thinking problem solvers we will need in the future. It can also provide tools for an important shift in the way schools are governed, a shift from rules and regulations set in some central office to local responsibility for meeting agreed-upon outcomes.

What I would like to do in conclusion is lay out a set of principles that could well guide the committee's thinking as it considers how to legislate for improved and equitable education throughout the Nation. These principles, taken together, comprise the elements of a new social compact for education that will ensure all students a fair shot at reaching high standards of achievement. I will mention only some. The full set is in the written testimony.

First of all, it is time to abolish the practice of expecting less from poor and minority children than from others. Student performance standards and the assessment systems that go with them should be designed to help bring all groups of students to the same high standards of performance.

Closely linked to that is the idea that we can no longer afford to educate a few to think and train many to do. The education standards set for the country for everyone should emphasize the ability to think and to use what they know outside a school setting. That means using complex, new forms of assessment.

Standards of performance need to be the same throughout the Nation, though States and districts can and should use different assessments to measure that performance. We are one country, not many, and no children should be denied the right to a first-class education because they go to school in one State rather than another, or in one school district rather than another. But, at the same time, we are a country of enormous diversity, and one of our challenges, a meetable one, is to design an assessment system that can recognize and honor all of those differences.

Standards need to be set through an open and public process that engages all sectors of society throughout the Nation. A really grassroots movement is needed, and needs to be organized and energized. Education standards are not the concern of educators alone. Parents especially, but also employers, child advocates and the public-at-large have a stake in what students learn, because this is the future of the country as well as the future of the individual students we're talking about.

New standards need to be deliberately tied to curriculum and designed as targets for instruction, not as neutral measuring devices. Teachers need to be actively involved in developing, grading and interpreting those assessments. Development and implementation of new assessments should be linked to professional development

and enriched curriculum so that all students have a fair opportunity to learn what is assessed.

Finally, the system needs to be designed so that it provides positive incentives for student effort. Demonstrated competence must lead to real opportunity for further education and jobs with advancement prospects.

Now, that's a lot, and there are skeptics who say it can't be done. They note that the new forms of assessment that we're calling for are untested, that they haven't yet proven themselves able to meet technical criteria of reliability, fairness and validity. Much of that criticism seems to imply that it would be better to stick with the discriminatory and educationally destructive current testing technology rather than invest in developing the new technical capacity we will need for the program I have outlined.

The cautions are, of course, appropriate. There is a lot of work ahead, and we are going to have to be careful about inappropriate uses of assessment results, especially if they're not accompanied by determined efforts to implement a full systemic reform program. But we can only build the new education system we need by getting started. We can't build it on paper. New forms of assessment and standard setting are a critical part of the process. There are risks to acting, but the risks of not acting are even greater. I urge that Congress move ahead with this program.

Thank you.

[The prepared statement of Dr. Lauren B. Resnick follows:]

Prepared Remarks of

Lauren B. Resnick, Director
Learning Research and Development Center
University of Pittsburgh
3939 O'Hara Street
Pittsburgh, Pennsylvania 15260
412-624-7485

before the

House Committee on Education and Labor
Subcommittee on Elementary, Secondary, and Vocational Education
Representative Dale E. Kildee, Chairman

February 4, 1992

I am Lauren B. Resnick, Director of the Learning Research and Development Center at the University of Pittsburgh. I appreciate the opportunity to speak with you today. I served as a member of the National Council on Education Standards and Testing and signed its recommendations document. This morning I would like to say why I favor the Council's recommendations and urge the Congress to move forward in adopting them.

I believe that the Council recommendations on education standards and assessments, joined with the provisions of the Neighborhood Schools Improvement Act (HR3320), provide a foundation for dramatic improvement in the nation's education system. There are three reasons, all closely linked: equity in educating all Americans, increasing the capacity of American schools to educate all students well, and creating a high performance education system that will blend national aspirations with local energy and decision making.

Education Standards and Equity

American children today do not have an equal chance to learn in school. Dramatic differences in resources available in rich and poor school districts persist and are growing. But the most profound inequities are often hidden from view; they persist even when children from different backgrounds attend the same schools. Even in such schools, expectations are not the same for all and the opportunities for learning are not equivalent. Almost everywhere, we expect less of children from poor and minority families; we therefore ask less of them; and, above all, we offer them less.

Accounts of these hidden, but profoundly disabling inequities abound. We treat children in ways that effectively limit what poor, minority and immigrant children have a chance to learn. Formal tracking systems aren't needed. All it takes is a first grade program in which the "high" reading groups are engaged in interpreting stories while the "low" reading groups are being drilled on recognizing words; in which "good" math learners solve interesting problems while "weak" math learners are practice computations they are not expected to understand. Carry this program of different *de facto* curriculum forward a few years and the initial, relatively small differences between children become large and unmanageable. Early differences in expectation have become a self-fulfilling prophecy.

For poor and minority parents, the differences in expectation, and therefore in the kind of instruction offered to their children, are often impossible to confront. Parents are told their elementary school children are doing fine, "working to their ability." How can they know that "working to ability" means working to a lower expectation than others? Or if they do suspect, what can they do to insist that higher expectations and more challenging instruction are appropriate for their children? It is not unusual to hear stories from poor or minority parents like the following: A child who was earning A's and B's in a predominantly minority school moved to a more mixed school in sixth grade. There he was judged "unable to read," given failing grades, and judged to be in need of remediation--to the astonishment and anger of his parents. Our schools today are operating on a double standard. With good intent--protecting children from undue stress and respecting individual differences--the education system is, often, effectively lying to poor parents and children. By the time families find out, it is too late--or at least much harder--to recoup.

Many poor and minority parents recognize that clear and public standards--the same ones for everyone--are the only way this situation is likely to end. Last Fall, the New Standards Project[1] ran a series of focus groups in several parts of the country, seeking information about the opinions of Americans of different social, economic and ethnic groups on the desirability of national education standards and assessments. The results showed support for national standards from all quarters. But the strongest support came from low income and minority parents who wanted their children to be judged by the same criteria as other children and who saw in standards a chance to improve the educational chances for their children.[2]

Increasing the Education System's Capacity

To take America confidently into the next century, preserving both our civic institutions and our standard of living, will require a much more highly educated workforce than we have today. We must now educate everyone to high standards of thinking, reasoning and practical application that in the past were reserved for an elite. Our schools are not organized for these new goals. The curriculum in place today in most schools is not designed to bring students to the high standards of attainment that their future and the nation's demand. Teachers and administrators are themselves mostly the product of a low-demand education system; they have not been educated for the new kinds of curriculum and teaching we now aspire to.

We are all imprisoned in our past. Only a massive bootstrapping operation can take us where we want to go. All resources need to be mobilized in coherent programs aimed at meeting ambitious educational goals. Bootstrapping ourselves into the future will require operating on all

---

[1]The New Standards Project is a privately funded consortium of 17 states and 6 school districts that is developing a system of shared standards and multiple assessments to those standards. New Standards, directed by Marc Tucker and myself, is a joint undertaking of the Learning Research and Development Center at the University of Pittsburgh and the National Center on Education and the Economy.

[2]A report of the focus group study, entitled "Public Response to the New Standards Project," by Dr. Vincent J. Breglio of Research/Strategy/Management Inc. is available from the National Center on Education and the Economy, 1341 G Street NW, Suite 1020, Washington, DC 20005.

February 4, 1992

fronts at once--standards, assessment, curriculum and teaching materials, professional development, preschool education, adequate health and social services. The Standards Council recognized that in its report, as does HR3320.

In addressing the desirability of education standards and assessments the Council was clear about why a new system of assessments is desirable. It is not to monitor education achievement and show yet again that we are doing badly--our National Assessment of Educational Progress (NAEP) does that very well. It is to provide a tool for raising the achievement levels of all children. To serve these ends the country needs new forms of assessment that will:

- exemplify targets of instruction, making it clear to everyone what kinds of capabilities students should be learning.
- be as rich and complex as the world for which students are preparing.
- be linked to local decisions about curriculum, so that they can legitimately organize teaching and learning.
- engage the loyalty and understanding of teachers and educators throughout the system.
- be reliable, fair and valid for their intended purposes.

Assessments that meet these demanding criteria can be major instruments in the bootstrapping reform of American education. They can change education practice in profound ways. With the richer, more complex assessments the Council envisaged, the kind that will be worth teaching and studying for because they exemplify true targets for learning, the natural desire of educators to prepare their students for the challenges ahead will work to raise achievement for all students.

That strategy can work, however, only if teachers and educators understand and are committed to the assessments as educational targets. Imposed from far away, without local participation in design, administration and grading, assessments--no matter how elegant--will not engage teachers' best efforts. They will not be able to effect the change in attitudes, habits and ways of thinking that are necessary for the new kind of education achievement set out in the national goals. That is why the Council was explicit in it call for *a varied system of state and local assessments linked to national standards*. Recent press coverage has misrepresented the Council's proposal. It did not call for a single national test. In fact, it was explicit in rejecting such a unitary test. Such a test would not serve the ends the Council had in mind. It would lock inequities into place and would not increase the education system's capacity

A High-Performance Education System

A system of national standards and locally developed assessments tied to them is the key to a high performance education system that can open real opportunities for all Americans. Such

8. Teachers should be actively involved in developing, grading and interpreting the assessments. Only in that way can assessments serve as a linchpin for improving instruction.

9. Development and implementation of new assessments should be linked to professional development and enriched curriculum so that all students have a fair opportunity to learn what is assessed.

10. The system must be designed so that it provides positive incentives for student effort. Demonstrated competence must lead to real opportunity for further education and jobs with advancement prospects.

There are skeptics who say this can't be done. They note that the new forms of assessment being called for are untested, that they haven't yet proven themselves able to meet technical criteria of reliability, fairness and validity. Such critics seem to imply that it would be better to stick with the discriminatory and educationally destructive current testing technology, rather than invest in developing the new technical capacity we will need for the program outlined here. Of course the cautions are appropriate. There is substantial work ahead, and we must be careful about inappropriate uses of assessment results, especially if they are not accompanied by determined efforts to implement a full systemic reform program. But we can only build the new education system we need by getting started, and new forms of assessment and standard setting are a critical part of the process. Education is in crisis, and failure to act is the biggest risk we can take.

February 4, 1992

a system can set achievement targets that are not just higher, but different in kind--able to produce the literate, thinking problem solvers we will need in the future. It can also provide the tools for an important shift in the way schools are governed, a shift from rules and regulations set at the center to local responsibility for meeting agreed upon outcomes.

To conclude, I would like to offer a set of principles that could well guide the Committee's thinking as it considers how to legislate for improved and equitable education throughout the nation. Taken together, they comprise the elements of a new *social compact for education* that will ensure all students a fair shot at reaching high standards of achievement.

1. It is possible to have an education system that is both equitable and excellent. Our system of standards and assessment can and should reflect that commitment. Its goal should be to raise the performance of virtually all students to world class standards, not simply to measure student performance. This will mean raising the top as well as closing the gap between the best and the worst student performances.

2. It is time to abolish the practice of expecting less from poor and minority children than from others. Student performance standards and assessment systems should be designed to help bring all groups of students to the same high standards of performance.

3. Standards of performance should be the same throughout the nation, though states and districts may use different assessments to measure student performance. We are one country, not many, and no children should be denied the right to a first-class education because they go to school in one state rather than another or in one school district rather than another. But, at the same time, we are a country of enormous diversity and the design of our assessment system must recognize and honor our differences.

4. The time is over when this country could afford to educate a few to think and train many to do. The education standards set for the country should emphasize the ability to think, and to use what they know outside a school setting. The assessments should use performance examinations, projects, exhibitions and portfolios to assess the capacity of students to apply what has been learned to the complex problems that they will encounter as citizens, family members, and members of the workforce.

5. Though all students should be held to the same standards, assessments should provide students with many ways to demonstrate their competence, enabling them to take advantage of the strengths of their particular cultural backgrounds and experiences. Assessments should be free of any cultural or racial bias.

6. Standards must be set through an open and public process that engages all sectors of society, throughout the nation. Education standards are not the concern of educators alone. Parents especially, but also employers, child advocates and the public at large have a stake in what students learn.

7. New assessments should be deliberately tied to curriculum and designed as targets for instruction and learning not as neutral measuring devices.

February 4, 1992

Chairman KILDEE. Thank you, Dr. Resnick.

Dr. Darling-Hammond.

Dr. DARLING-HAMMOND. Thank you.

I don't need to tell you that over the last couple of years we've seen an array of proposals at the Federal level to cure America's educational problems with national tests or testing systems. These have included proposals for an individual NAEP, proposals for Educate America, the American Achievement Test proposed, the America 2000 plan, and now a national system of assessments—this last, in many respects, the most responsible and well considered of these proposals, but one that should still be subjected to serious scrutiny.

I support many of the intentions and proposals outlined in the report that's been described already, but I think it's also important to raise some questions to guide the deliberations of the Congress on this. Many ideas and assumptions underlie the national testing proposals. One is that other countries are beating us educationally because of their national tests. This is an interesting perception because many of the countries that are performing substantially better than the United States on international assessments do not, in fact, have national tests. Germany, Australia and Canada, for example, have State or provincial systems of education and assessment. Japan, contrary to the assumptions of many, has no national testing system.

Why so much emphasis on national tests? Some would say it's an attempt to appear to be improving education without having to invest dollars or energy to change schooling. There may be something to that argument and that's why I think it's important that there be teeth in the delivery standards idea that is proposed in the NCEST report.

Some claim, rightly I believe, that most current U.S. tests are counterproductive to the development of higher level thinking and performance skills. There is a legitimate set of questions, though, as to whether and how national standards will improve on the many innovative and important efforts already underway in States and local districts, or whether they will constrain and distort them.

There are important issues. For example, consensus on values. How will we handle multicultural perspectives on a national level? As well as on the nature of learning desired. Will we end up with a compromise between the standards of the most high achieving States and the least high achieving that leaves us somewhere in the middle? Also, what is the kind of learning we want to encourage? Do we want to aim at the problem-oriented skills used in interdisciplinary tasks, as was described in the SCANS report, or the kinds of tests and traditional—and some would argue—rigid academic disciplines that are proposed to start the new testing system? And should that be determined at the national level?

I don't pretend there are not appropriate answers to those questions, but I think they need to be asked and examined as we proceed to figure out what approaches to assessment reform makes sense.

Where is the line between a national curriculum and the national specifications of curriculum standards and goals that are called for? How does one reconcile demands for comparability in test

scores with promises to respect diversity, innovation and local democratic processes?

Finally, what use will tests be put to and how will those help and hurt American students and the American school system? Americans have a tradition for misusing tests for inappropriate decisions in ways that research confirms have harmed educational quality and have seriously harmed educational equity.

These concerns have been raised by many over recent months. Nearly a year ago, the American Psychological Association issued a statement raising a series of these concerns and stated that a high-stakes national testing system would be counterproductive. Last June, the APA, along with the American Educational Research Association and the National Council on Measurement and Education, issued a joint statement based on testing research, urging the Nation to slow down and think differently about test-based accountability.

Many of the individuals most deeply involved in the most serious school reforms in this Nation have raised concerns about the rush to a national testing solution. And civil rights and child advocacy organizations have asked questions about the potential benefits and harm to students that have not yet been answered.

As a consequence of these concerns and conversations among many of us involved in educational reform, a statement laying out criteria for evaluating the plethora of testing proposals that have emerged was drafted 2 weeks ago, just before the release of the NCEST report and before we knew what the contents would be. The statement is aimed at informing Members of Congress, as well as the public and press, of our concerns regarding national testing proposals.

The concerns are: one, preserving local involvement in standard setting and assessment development; two, ensuring that tests are not used to administer high-stakes consequences to students and schools; and three, calling attention to equity concerns in the availability of educational resources.

The 53 individuals who signed during the three or four days in which it was circulated represent leadership in the educational research and testing communities, including the current president and numerous past presidents of the American Educational Research Association, the National Council for Measurement in Education, the National Academy of Education, and organizations such as the Educational Testing Service and the College Board. Individuals deeply involved in school restructuring, including Ted Sizer of the Coalition of Essential Schools, James Comer of the School Development Program, John Goodlad of the Center for Educational Renewal, and many others, and individuals representing child advocacy and civil rights organizations, including the Children's Defense Fund, the National Urban Coalition, the National Urban League, the NAACP Legal Defense and Education Fund, and the Quality Education for Minorities Project. Of the signers include members of the NCEST group and its task forces.

I would like to enter into the record, to shape the considerations of the panel as it proceeds. It says, in the current ferment around school reform, considerable emphasis has been given to the development of a new national test of educational achievement as a

means of establishing higher standards for students and holding schools accountable. We believe that the development and use of such a test is not in the best interests of our educational system or our students. While we all agree that dramatically higher educational standards are needed for American schools, we believe that the pursuit of such standards does not require and could be severely compromised by such a national examination.

Along with many others, we believe that more challenging curriculum goals in American schools must be supported by more valid performance-based assessments of what students understand and know how to do. We are not opposed to all approaches to assessment reform, but we believe there are important criteria for determining the merits of various proposals. For assessment to be a constructive agent in teaching, learning, and school improvement, we believe the following conditions are necessary. We are also persuaded that some approaches could be seriously counterproductive.

First, we believe that the primary goals of an assessment system should be to improve teaching, support learning, and increase educational opportunity, not to become simplistic arbiters of rewards or sanctions for students or schools. Members of the measurement community agree that the limited information available from test results should not be used as sole determinants of students' educational futures—for example, for tracking, promotion, or graduation decisions—or for assigning consequences to schools. Past research and experience have shown that these inappropriate uses of test results ultimately narrow curriculum, thus lowering standards, and create perverse incentives for schools to artificially raise scores—and this is key when we think about an outcomes-based approach to accountability—by pushing out low-scoring students rather than improving the quality of education. Imposing national standards through a high stakes assessment system that promotes such inappropriate uses of test results would harm teaching and learning and further reduce educational opportunities for disadvantaged students.

Second, we believe that in order to achieve challenging educational goals, local educators must participate in developing and implementing assessment practices that express the considered values and standards of their school communities and are clear and open about the capabilities and performances desired of students.

Unless local communities are integrally involved in standard setting, they will not own the standards, nor will they develop the will and capacity required for ongoing school investment and improvement. America's commitments to intellectual freedom and democratic decision making require that national curriculum ideas be open to rich examination and imaginative local revision.

We want to see an accountability system rooted in local efforts to take stock, supported and monitored by districts and States, and informed, but not controlled, by national perspectives on achievement. Such a system recognizes that the creation of rigorous standards does not require standardization.

Finally, we believe that any real effort to create accountability in American schools must focus equal or more attention on improving the capabilities of children to learn and schools to teach as it does on gaging educational outcomes. Parents and taxpayers have

a right to expect that accountability policies will aim to provide good education, rather than merely measuring schools.

Given the tremendous differences between today's achievements and the goals set for America 2000, the inadequate supports for children and families in American society, and the dramatic inequalities among schools' resources, any policy to establish benchmarks for achievement without creating equity in the educational resources available to children, would be a cruel hoax.

It is important to note that there the NCEST report does acknowledge two of these concerns. For investments in educational equity, the school delivery standards and system assessment standards are very important, and for local development and implementation. However, there's a point of tension between the views of the educators and child advocates expressed in this statement. In the NCEST report on the question of States, the report claims that the assessments can be used for such purposes as high school graduation, college admission, continuing education, certification for employment, system accountability and rewards and sanctions to teachers and schools.

It is important also to note that no other country uses its assessment system for this array of high stakes purposes, and none for rewards and sanctions to teachers and schools. There are good reasons for this. There are a lot of reasons why this could be problematic. First of all, high stakes can corrupt assessments—and I believe that Walt Haney will address that in more detail, so I won't spend any time on it. But most important for the equity concerns is the fact that no test is adequate, in and of itself, to base a major decision regarding students' future options. No technology, no matter how glamorous, can capture the important and diverse qualities and potentials of human beings. That's a point that has been raised time and again by the measurement community.

Cutting off options is a serious thing to do. There is a difference in the ways in which we currently use such assessments, as the Advanced Placement Test referred to earlier, for example, in Jaime Escalante's classroom shown in "Stand and Deliver," the SAT test, and the nature of States which, in fact, would not just give a score or information but would cut off students' future options in life.

Let's say that high, universal, world-class standards might be equivalent to something like 1400 on the SAT. How many of us, if we didn't score—I won't ask.

[Laughter.]

How many of us should have had our future options for high school graduation and employment cut off based on a standard that is used for that kind of purpose?

It's one thing to have an assessment system that delivers information and scores that can be factored into decisions and school improvement and instruction improvement and, in fact, decisions that colleges or others would make. It's another to create a system which, in fact, says, based on this one set of information, we make the decision that will allow you to further develop your talents in society, or to be cut off from developing them. So I think we have to be very careful when we think about what the uses of an assessment system will be—for all of us.

52-840 O - 92 - 2

Of course, there are special concerns for the most underserved students in our society in that regard. My daughter put it this way. She said the problem with the standard-setting stuff is that it tends to penalize those who are different. Now, if a 15 year old can acknowledge that, I think most of us can understand it. Somebody makes decisions at some point about what is valued.

The beauty of an assessment system of the sort described in this report is that it might be a very full and rich way, with much higher standards for people to guide, instruct, and understand students. The danger is that inappropriate high stakes are attached to it. We may cut off the opportunity for our Nation to develop the talent, which we can't ever fully enumerate and quantify, of some of its citizens to go on and develop further in life.

Finally, stakes for schools create a perverse incentive—we have a lot of research on this now—to dissuade talented staff from teaching in difficult schools and to push out or keep out low achieving students. If schools are evaluated on the proportion of students who achieve well on tests, they will, if they have the option, keep out those students or push out those who do poorly on the tests as a way of boosting the measure on which they're evaluated. If we care about encouraging equity, we should be applying accountability to States and school districts in their treatment of students as the primary way in which we use assessments.

The assessment task force and NCEST acknowledged these issues and raised many questions, and I urge you to attend to those.

I would be much more supportive of a voluntary national assessment system if it focused explicitly on educational information and improvement, if it were clearly tied to legislation to invest in school development, especially for under-financed schools, and if it eschewed high stakes intentions that hold the greatest danger for destroying both the integrity of the assessments and the appropriateness of incentives to schools and students.

[The prepared statement of Dr. Linda Darling-Hammond follows:]

Submitted as part of the testimony of Linda Darling-Hammond

Response to National Testing Proposals

In the current ferment around school reform, considerable emphasis has been given to the development of a new national test of educational achievement as a means of establishing higher standards for students and holding schools accountable. We believe that the development and use of such a test is not in the best interests of our educational system or our students. While we all agree that dramatically higher educational standards are needed for American schools, we believe that the pursuit of such standards does not require -- and could be severely compromised by -- such a national examination.

Along with many others, we believe that more challenging curriculum goals in American schools must be supported by more valid, performance-based assessments of what students understand and know how to do. We are not opposed to all approaches to assessment reform, but we believe there are important criteria for determining the merits of various proposals. For assessment to be a constructive agent in teaching, learning, and school improvement, we believe the following conditions are necessary. We are also persuaded that some approaches would be seriously counterproductive.

First, we believe that the primary goals of an assessment system should be to improve teaching, support learning, and increase educational opportunity, not to become simplistic arbiters of rewards or sanctions for students or schools. Members of the measurement community agree that the limited information available from test results should not be used as sole determinants of students' educational futures -- e.g. for tracking, promotion, or graduation decisions -- or for assigning consequences to schools. Past research and experience have shown that these inappropriate uses of test results ultimately narrow curriculum, thus lowering standards, and create perverse incentives for schools to artificially raise scores by pushing out low-scoring students rather than improving the quality of education. Imposing national standards through a high stakes assessment system that promotes such inappropriate uses of test results would harm teaching and learning and further reduce educational opportunities for disadvantaged students.

Second, we believe that in order to achieve challenging new educational goals, local educators must participate in developing and implementing assessment practices that express the considered values and standards of their school communities and are clear and open about the capabilities and performances desired of students. Unless local communities are integrally involved in standard-setting, they will not "own" the standards, nor will they develop the will and capacity required for ongoing school investment and improvement. America's commitments to intellectual freedom and democratic decisionmaking require that national curriculum ideas be open to rich examination and imaginative local revision.

We want to see an accountability system rooted in local efforts to take stock, supported and monitored by districts and states, and informed -- but not controlled -- by national perspectives on achievement. Such a system recognizes that the creation of rigorous standards does not require standardization. It would extend beyond examinations to include cumulative evidence of student accomplishments and of school practices, including inspection or audit systems that require each school to examine its own goals and performance, disclose what its indicators of achievement reveal, and report to its community, striving for continual improvement.

Finally, we believe that any real effort to create accountability in American schools must focus equal or more attention on improving the capabilities of children to learn and schools to teach as it does on gauging educational "outcomes." Parents and taxpayers have a right to expect that accountability policies will aim to provide good education, rather than merely measuring schools. Given the tremendous differences between today's achievements and the goals set for America 2000, the inadequate supports for children and families in American society, and the dramatic inequalities among schools' resources, any policy to establish benchmarks for achievement without creating equity in the educational resources available to children would be a cruel hoax.

The attached statement is aimed at informing members of the Congress as well as the public and press of our concerns regarding national testing proposals. The concerns are for 1) preserving local involvement in standard-setting and assessment development, 2) ensuring that tests are not used to administer high stakes consequences to students and schools, and 3) calling attention to equity concerns in the availability of educational resources.

Initial signers include:
Gregory R. Anrig, President, Educational Testing Service
Sue E. Berryman, Director, Institute on Education and the Economy, Teachers College, Columbia University
Julius Chambers, Director-Counsel, NAACP Legal Defense and Education Fund
David K. Cohen, John A. Hannah Professor of Education, Michigan State University
James P. Comer, M.D., Maurice Falk Professor of Child Psychiatry, Yale Child Study Center; Associate Dean, Yale School of Medicine, Yale University
Eric Cooper, Executive Director, National Urban Alliance for Effective Education
Larry Cuban, Professor of Education, Stanford University Past President, American Educational Research Association
Linda Darling-Hammond, Co-Director, National Center for Restructuring Education, Schools and Teaching and Professor,Teachers College, Columbia University
Ramona Edelin, President, National Urban Coalition
Marian Wright Edelman, President, Children's Defense Fund
Elliot Eisner, Professor of Education, Stanford University; President Elect, American Educational Research Association
Michelle Fine, Goldie Anna Professor of Psychology in Education, University of Pennsylvania
Howard Gardner, Co-Director, Project Zero, Harvard Graduate School of Education
Keith Gieger, President, National Education Association, National Center for Innovation
Bernard R. Gifford, Vice President of Education, Apple Computer; Founding Chair, National Commission on Testing and Public Policy
John I. Goodlad, Professor and Director, Center for Educational Renewal, University of Washington Seattle
Edmund Gordon, Emeritus Professor, Yale University
Gary A. Griffin, Professor of Education, University of Arizona
Walt Haney, Senior Research Associate, Center for the Study of Testing, Evaluation & Educational Policy, Boston College
George H. Hanford, President Emeritus, College Board
Delwyn Harnisch, Associate Professor, University of Illinois Urbana-Champaign
Harold Howe II, Senior Lecturer Emeritus, Harvard Graduate School of Education

Philip W. Jackson, Professor of Education and Psychology, University of Chicago; Past President, American Educational Research Association

Richard Jaeger, Professor of Educational Research Methodology and Director of the Center for Educational Research and Evaluation, University of North Carolina

Sylvia Johnson, Acting Editor-in-Chief, Journal of Negro Education; Professor of Research Methodology and Statistics, Howard University

Judith Lanier, Dean, Michigan State University College of Education; President, Holmes Group

Ann Lieberman, Co-Director, National Center for Restructuring Education, Schools and Teaching and Professor,Teachers College, Columbia University; President, American Educational Research Association

Shirley McBay, President, Quality Education for Minorities

Milbrey McLaughlin, Professor of Education and Director, National Center on Secondary School Context, Stanford University

George Madaus, Director, Center for Testing, Education, Evaluation and Public Policy and Boisi Professor of Education and Public Policy, Boston College

Deborah Meier, President, Center for Collaborative Education and Director, Central Park East Secondary School, New York City, New York

Joe Nathan, Director, Center for School Change

Jeannie Oakes, Professor, University of California Los Angeles

Gary Orfield, Professor of Education and Social Policy, Harvard Graduate School of Education

Vito Perrone, Director, Teacher Education Programs, Harvard Graduate School of Education

Robert Peterkin, Harvard Urban Superintendents Program and National Urban Alliance

Barbara Plake, President, National Council for Measurement in Education; Professor of Psychology and Measurement, University of Nebraska

Philip C. Schlechty, President, Center on Leadership and School Reform, Louisville, Kentucky

Donna Shalala, Chancellor, University of Wisconsin

Richard J. Shavelson, Dean and Professor, University of California Santa Barbara

Lorrie A. Shepard, Professor of Education, University of Colorado at Boulder

Lee Shulman, Professor of Education, Stanford University; President, National Academy of Education

Ted Sizer, Chairman, Coalition of Essential Schools, Brown University

Marshall S. Smith, Dean, School of Education, Stanford University

Robert Stake, Professor, University of Illinois Urbana-Champaign

[illegible] Stewart, President, College Board

Gary Sykes, Associate Professor, Departments of Educational Administration and Teacher Education, College of Education, Michigan State University

Concepcion Valadez, Associate Professor, Graduate School of Education, University of California Los Angeles

Michael Webb, Director of Education and Career Development, National Urban League
Grant Wiggins, Director of Programs and Research, Center on Learning, Assessment and School Structure
Eliot Wigginton, Foxfire Network, Rabun Gap, Georgia
William Julius Wilson, University Professor of Sociology and Public Policy, University of Chicago
Arthur Wise, President, National Council for Accreditation of Teacher Education

Persons listed above have signed this statement as individuals. Organizational affiliations are listed for the purpose of identification only.

THE NEW YORK TIMES EDUCATION WEDNESDAY, JANUARY 29, 1992 B9

Prominent Educators Oppose National Tests

By SUSAN CHIRA

Intensifying the debate over national tests for schoolchildren, more than 50 prominent educators and testing experts have released a statement opposing such examinations just as a major educational advisory panel has endorsed them.

Last Friday the advisory panel, the National Council on Education Standards and Testing, which is composed of both political and educational leaders, made public a final report calling for national curriculum standards and national tests as the best way to insure that all schools improve. This would be a radical departure from the longstanding American tradition of local control of curriculums that has led to a patchwork of uneven course requirements, tests and school results.

Details of the plan were reported in mid-December, when the council finished its draft of the report.

Penalty for Students Cited

The report is dividing the education community. The 50 educators opposing national tests said in their opposition statement last week that such examinations would end up penalizing students who attend bad schools and depriving local communities of control over what their students learn.

In the past, when schools have been rewarded or punished on the basis of test scores, the educators said, research studies have shown that many schools tried to raise scores artificially by excluding low-scoring students.

Measuring schools by test results is unfair, they argue, because there are huge gaps in how much money different schools receive and in the kinds of students they serve. "Any policy to establish benchmarks for achievement without creating equity in the educational resources available to children would be a cruel hoax," the statement said.

Signers include James P. Comer, the Yale psychiatrist whose program to reform schools is widely hailed as a national model; Theodore R. Sizer, a professor at Brown University who heads a network of more than 200 schools committed to improving students' performance, and Marian Wright Edelman, president of the Children's Defense Fund. The statement was also signed by two members of the standards and testing council who expressed reservations about the the council's endorsement of national testing. They are Keith Geiger, president of the National Education Association, the country's largest teachers' union, and Marshall S. Smith, dean of Stanford University's School of Education.

Chairman KILDEE. Thank you very much.

Dr. Haney.

Dr. HANEY. Thank you, Mr. Chairman. I am pleased to be here before the committee today to offer commentary on the NCEST report, entitled "Raising Standards for American Education."

I won't try to detail my professional background, except to say that I have been working in the area of educational testing and evaluation for close to 20 years. I am affiliated with Boston College in several capacities. But I should emphasize that my comments here this morning represent my own thoughts on this recently released report.

I would mention, however, that I had the privilege of testifying before this subcommittee about 18 months ago concerning the report of the National Commission on Testing and Public Policy, which I won't say more about further, except to emphasize that the recommendations of that national council are well worth revisiting in light of the recommendations of the NCEST council.

Let me mention that in my prepared remarks I indicate there are a number of aspects of the NCEST report that I personally do like. For example, at least the report raises the point that, when it comes to educational standards, we need to be concerned with standards not just for students but also standards for schools and school systems. What is odd, though, is why the standards for students should be national but the standards for schools should be up to the States. There's no rationale provided in the report for that distinction. Moreover, it seems pretty obvious to me, for both constitutional reasons and historical reasons of educational governance, quite the reverse set of priorities would be more appropriate than the priorities set out in the Council report. It's a mystery why they recommended that because it's just not explained.

There are a number of other things that I liked in the report but, to be honest, I won't take the time to sort of point out areas in which I concur with things in the report because, overall, my reactions to the Council report was one of great disappointment. Let me try to summarize my concerns under four rubrics—and I could say that, after hearing Dr. Resnick and Dean Smith talk this morning, I remain a little confused about what the real agenda behind this system is. But I won't react to their comments now but will save those until perhaps some time for exchange.

First of all, it seems to me that the Council has failed almost completely to meet its charge of assessing the desirability of a national system of standards and assessments for all schools in the Nation. Basically, my point here is simply that we cannot evaluate the desirability of a treatment unless we're given a diagnosis. It seems to me the Council completely failed to undertake any kind of wide-ranging diagnosis of what is fundamentally wrong with the American system of education.

There are many examples of competing needs in education in our society, ranging from recruiting better teachers to providing day care. On many of such indicators, our Nation lags far behind many other nations in the world. How are we to decide, on the basis of the Council report, whether it's more desirable to have some national system of standards and assessments or to invest in immuni-

zation for children, another social indicator, by the way, in which we lag far behind many other nations.

I would note that the absence of a broad consideration of the desirability of such issues is particularly notable in the NCEST report in light of the fact that goal one of the national education goals for the year 2000 is that "all children will start school ready to learn." Why, since the Council claims to be developing a system of assessments to assess progress toward the national goals, was there no attention given to the very first of these national education goals?

More generally, however, it seems to me the fundamental ambiguity about the Council recommendation is that the priority purposes of national assessments and the priority focus for national assessments is simply not clarified. I note that the Council seems to carry through the subject matter focus set out in the national education goals emphasizing English, math, science, history and geography. There's nothing wrong with those subjects. But stop a minute. Are those the subjects that should be the focus of a national educational system that is intended to drive instruction and learning? Let me encourage you to think about that by asking a simple question.

Do you remember when President Bush released the national education goals a couple of years ago? He said at the time that he was himself going to return to being an adult learner and he was going to study something that he really wanted to learn. Do any of you happen to remember what the subject was that he chose to study? Was that a priority subject in the national education goals? Nope. Not mentioned.

Well, I would urge you to reflect why President Bush might have chosen to study that subject rather than one of the core academic subjects, because I think the reasons for his decision probably say a lot about many people's learning, not just that of the President. I would note—and I've provided documentation in my written testimony—that public opinion surveys indicate little concurrence with the subject matter priorities identified in the national goals and carried through in the NCEST report.

My purpose in raising this issue is not to try to quibble about whether geography should get more attention history, or whether computers should be taught in schools instead of geography. Rather, my point in raising this issue is that I think in the NCEST report there is a pretension about widespread consensus for the priority goals of American education that does not exist. Once we scratch beneath the surface of rhetoric about national standards, we see that there is far less consensus than the Council pretends with regard, for example, to the relative priority to be placed on not just different subjects but on the balance of educational aims across the grade levels, on the appropriate balance between academic learning versus vocationally-oriented learning, and in particular, concerning the relative emphasis that ought to be placed on what is alluded to toward the tail end of goal three as "citizenship skills."

I note that national polls indicate that when people are asked to identify the major problems in schools, what they for many years have tended to identify are a lack of discipline and use of drugs,

not lack of academic learning. The reason I say this is that it seems to me the Council ignores the point that there are many roles of our schools that cannot be reduced to external objective assessment, what might be called "civic virtues" or even—though it's an archaic term—"character development."

With regard to purposes of assessment, I note that on page 26 of the Council report there are five different purposes for the national assessment system listed. But if you look closely at those five purposes, they really encompass about a dozen different possible functions. My chief concern with the Council report is that you can't have one assessment, or even one comparable system of assessment, that can equally well serve all of those purposes. If we try to focus an accountability system on accountability for schools and teachers, we know, from both contemporary and historical experience, that that accountability function will tend to corrupt assessments so that scores may appear to go up but the learning of students will not go up in tandem with the scores on the assessments. This is true not just of multiple choice tests but performance tests and essay tests, we know from both contemporary and historical experience.

I believe that the Council also gave inadequate attention to the feasibility issues, the technical issues of assessment development. There are many things that could be said about that. I will not say a great deal, since time is limited. But let me try to give one concrete example.

The Council said that assessment results need to be comparable in order to be useful. At the same time, the Council recommends that the assessments ought to be voluntary. However, it is absolutely clear from the research over the last decade that if test results are compared from one school or school system that regularly uses the test, the results do not mean the same thing as in another school which does not regularly use the same test. To put the matter even more simply, you can give the AP calculus test in a couple of different schools, but if one school has organized its calculus course around the AP calculus test, the results don't mean the same thing as in a school that has organized its calculus course around the history of mathematical thinking.

Basically, I don't want to say more right now, except to express my overarching general criticism of the Council report; that is, it presents no clear or consistent vision of how new assessments will help to improve teaching and learning in the schools. I comment in more detail on specific recommendations in my written testimony, so I won't recap my written testimony here.

But what I would like to say is that this general theory that has been espoused vaguely in the Council report, more concretely by Dean Smith and by Professor Resnick, about national assessments and equity, I hear that vague theory very often, that high national standards are going to motivate schools and teachers to bring all students to meet these high standards, that they're going to motivate teachers and schools. Well, we could talk about many of the anecdotes that are brought forward to advance that theory, such as "Stand and Deliver." But what I would like to point out is that graduation from high school in our country has had attached to it many important social sanctions for many decades. If you want to

get a job, other than menial labor, you have got to have a high school degree. If you want to get into college, you have got to have a high school diploma. If you, indeed, want to join the military, you really have to have a high school diploma. You can get in if you don't have a high school diploma, but then your test scores have to be high.

Despite all those sanctions attached to high school graduation, we still in this Nation have approximately 25 percent of our youth not completing high school. So I would ask Dean Smith and Dr. Resnick, if longstanding sanctions attached to high school graduation—which is much easier to achieve that some foreign high stakes test—if high school graduation, which already has attached to it many more sanctions than are proposed for this new national assessment, do not motivate kids just to stay in school, how can we possibly think that a key to motivating youngsters is some external test of unproven validity? I just don't buy that argument.

I will close for now because I know time is short, and for specific comments I would refer you to my written testimony.

[The prepared statement of Dr. Walter M. Haney follows:]

Testimony of Dr. Walter M. Haney
before the
Subcommittee on Elementary, Secondary and Vocational Education

House of Representatives
U.S. Congress

February 4, 1992

Washington D.C.

Walter M. Haney

Senior Research Associate

Center for the Study of Testing,

Evaluation and Educational Policy,

319 Campion Hall

Boston College

Chestnut Hill, MA 02167

ph. 617-552-4199

EMail: Haney@BCVMS (Bitnet)

Haney@bcvms.bc.edu (Internet)

Haney (Applelink)

Haney Testimony, 2/4/92, p. 2.

I am pleased to appear before this Committee today in order to offer commentary on the report released January 24 by the National Council on Education Standards and Testing, entitled *Raising Standards for American Education*. Let me start by briefly mentioning who I am, and then after explaining that while I like some aspects of the report, summarize why I think the report overall is unfortunately superficial and unlikely to contribute much toward improvement of education in the United States.

First off I should mention that I am currently a senior research associate in the Center for the Study of Testing, Evaluation and Educational Policy at Boston College, where I am also an associate professor and teacher in the School of Education. I have worked in the field of educational evaluation and testing for close to 20 years, but am also professionally and personally interested in the potential of information technology to improve instruction and learning. In case my professional credentials would be of interest to you, I have brought along a copy of my resume, but I should mention that my interests in improving education are not restricted solely to my professional life but arise also in various volunteer activities, including serving for some years on my local school board. Also worth mentioning is that I had the privilege of testifying before the House Subcommittee on Elementary, Secondary and Vocational Education in June 1990, concerning the work of the National Commission on Testing and Public Policy -- and whose work and recommendations, I might add, are well worth revisiting in light of the NCEST report.

Some Plaudits.

I might start by saying that I was heartened a few things in the NCEST report. The report at least raises, for example the point that when it comes to educational standards, we need to be concerned with standards not just for students but also for schools and school systems. Though why standards for students should be national but those for schools should be up to the states is certainly a mystery to me. Indeed, given the historical role of national government in education, just the reverse would seem far more appropriate. Also, the Council shows insight in proposing not just one test or set of tests, but a two tiered assessment, with one set of assessments aimed at individuals and the other aimed at gauging the educational health of the nation. The Council also at least mentions that time and research will be required into the validity and fairness of new assessments before they could reasonably be used for making "high-stakes" decisions about individuals, for example with regard to college admissions or employment. Finally, I note with approval that the Council recommends that the original guiding purpose of the National Assessment of Educational progress be maintained and that NAEP not be used to try to produce tests for individual students -- a move which in my view would have had great potential for undermining the useful role that NAEP has played for close to a quarter of a century in gauging the general levels of learning of young people in the United States.

Major Failures of the Council Report

Despite a few positive reactions to the Council report, overall, the Council report is in my view a disappointment, and promises to contribute little toward improving education and learning in the United States. There are many reasons for this conclusion, but since time is limited, here let me mention just four major concerns:

- Failure to assess desirability;
- Failure to clarify priority purposes and focus of assessments;
- Failure to assess feasibility of valid and comparable assessments;
- Lack of clear vision of how assessments would lead to increased learning.

1. Failure to assess the desirability of a national standards and assessments. The Council has failed almost completely to meet its charge of assessing the desirability of a national system of standards and assessments. Any reasonable assessment of the desirability of such a national standards-assessment system ought to be based in a consideration of the state of education in the United States and an analysis of the problems of education in America. Basically my point here is we simply cannot evaluate the desirability of a treatment until we have a clear diagnosis. Since NCEST report provides no clear diagnosis of the state of education in the United States, it does not provide any reasoned basis for the desirability of new national assessments. Thus, the Council appears to have completely failed to consider whether a new system of standards and assessments is desirable in comparison to a host of other obvious needs in American education such as developing more equitable means of school financing, better maternal and infant health care systems, full-funding of Head Start, or recruiting better teachers. Thirty-seven pages into the Council's 41 page report, mention is made of needs for developing policies and practices regarding curriculum, professional development and community and family supports. However, the Council seems to hold that such matters are the province of states and localities and that at the national level we should address mainly education outcomes, rather than inputs or processes. However, from a policy point of view, when we know that out nation lags behind many other nations in indicators such as infant mortality, immunization of children, and preschool care, and when we know that early childhood development determines much about learning later in life, it is vital to ask whether a new national system of educational assessment is desirable relative to investing more in building such fundamental foundations for learning. The absence of such a broad consideration of desirability is particularly notable in light of the fact that Goal 1 of the National Education Goals for the year 2000 is that "all children will start school ready to learn." Why, I would ask the Council, since you addressed the need for a national system of educational assessment to monitor progress toward the National Education Goals, did you not give serious attention assessing progress toward the very first national education goal?

2. Priority Purposes and Focus of National Assessments

The answer which Council members might well give to my question is that the Council was asked mainly to address the Goals 3 and 4 of the national education goals. But even here, I find that the Council report is superficial and unthoughtful with regard to both the priority purposes and focus of national education standards.

Regarding focus, I note the apparent assumption of the Council that priority subject matter focus of the assessment system ought to be English, math, science, history, and geography. But why these subjects -- other than the fact that they were identified in the hastily developed national goals for education? Let me illustrate my concern with the subject matter focus of Council proposals in two ways. First, if you recall when President Bush announced the America 2000 plan, he mentioned that he himself was returning to the role of adult learner, to study something. Do you recall what it was? Was his chosen subject of study one of the five subjects mentioned in the National Goals? No it was not. It is worth considering why not.

Second public opinion provides little support for the subjects the Council tells us should be the focus of new national assessments. In the 1990 Gallup poll of the public's attitudes toward the public schools, respondents were asked which subjects should be given more emphasis than now received in high school. Percentages of the national sample responding "more emphasis" were as follows (Elam, 1990, p. 49):

Math	80
English	79
Computer training	79
Career education	73
Science	68
History/US Gov't	65
Vocational education	65
Health education	62
Business	60
Geography	53

My purpose in raising questions about the subject matter focus recommended by NCEST is simply to warn about what I believe to be a fallacious assumption undergirding the Council report and many recent discussions about national tests, namely the presumption that there is widespread consensus about the priority goals of American education. Once we scratch beneath the surface of rhetoric about national standards, we see that there is far less consensus than the Council pretends, with regard for example, to the relative priority to be placed on different subjects of study, on the balance of educational aims across the grade levels and on learning that is more academically or more vocationally oriented. Moreover, there are important issues to be addressed, but which are swept under the rug in the Council report about the the role of schools in the United States in promoting

cognitive learning versus their role in promoting what is referred to toward the tail end of Goal 3 as citizenship skills. Note for example that when asked to identify the biggest problems facing local public schools in 1990, a national sample of respondents identified "use of drugs," lack of discipline," and "lack of proper financial support" as the top three problems (Elam, 1990, p. 53). In short there is ample evidence, none apparently considered by the Council, that what might be called "civic virtues" or even "character development" ought to be higher priority on our national educational agenda than suggested by the Council.

With regard to purposes of assessment, the Council report is similarly vague and superficial. Though the main impetus for new national tests, as indicated in the *America 2000* , was for greater educational accountability, the Council report now lists five different purposes, which if examined closely, actually encompass more than a dozen quite different purposes.

The problem with the Council report in this regard is twofold. First of all from both technical and practical points of view, the fundamentally different assessment purposes outlined in the Council report require different kinds of assessments and different kinds of validation. In this regard, the full Council report largely ignores the advice of its Assessment Task Force that no one assessment can serve such a range of purposes. To help illustrate this fundamental point, I append to my testimony, an analysis of the need for fitting assessments to particular functions.

Given the range of purposes the Council report outlines for new national assessments, a major failing is that the Council gives no clear recommendation as to the priority purposes toward which new national assessments ought to be aimed. In a moment I will elaborate on my own sense of such priorities, but at a minimum, I argue most strongly that until clear priorities as to the purposes to be served by new national assessments are set out, it would be foolish to waste more time or resources in setting up a structure to oversee the development of new national assessments.

3. Inadequate consideration of feasibility. The fundamental problem here is that technical consideration of characteristics of assessment such as reliability and validity must be addressed with regard to the purpose of assessment. Since the Council has not been clear as to the purpose served by the proposed national assessment system, one cannot adequately assess such issues an validity and reliability. One could discuss such technical issues at great length, but to illustrate, let me provide one simple example. The Council says that assessment results need to be comparable in order to be useful. At the same time, it maintains that the assessments must be voluntary, to be used by states as they choose. But if one thing about school testing became clear over the last decade, it is that test results from one school that regularly uses one kind of test are simply not comparable with those of other schools that do not. To put the matter more simply, an AP calculus test may be given in two schools, but if one school has

organized its calculus course around the AP calculus test, while the other has focused its calculus program around the historical development of mathematics, the results will not mean the same thing.

4. No clear vision of how national assessments will lead to increased learning. My fourth and overarching general criticism of the Council report is that it presents no clear or consistent vision of how new assessments will help to improve teaching and learning. Implicitly, the Council report seems to be relying on several different -- and somewhat contradictory -- ideas as to how assessment will improve learning. This is most unfortunate because until one articulate a clear vision of how a proposed assessment is to work, it is impossible to evaluate its desirability, its feasibility or its likely validity for its intended purpose.

****

Rather than critiquing the Council report further, let me briefly share with you what my own sense of the priorities for new assessments ought to be, and some very specific recommendations.

1. First off, if the prime motivation for new national assessments is for greater educational accountability at national and state levels, then the we should forget completely the first component proposed by the Council (individual student assessments) and focus entirely on the second component, namely "assessments of representative samples of students from which inferences about the quality of programs or educational systems could be made" (p. 26).

2. Second, my own view is that we are in far less need of more assessment for accountability or for management purposes by state or national educational policy-makers than we are in need of assessments that help individual students learn and individual teachers and parents help them learn. Thus I maintain that if there are a variety of purposes that <u>might</u> be served by new national assessments, then the #1 aim ought to be that they help individuals learn. In the paper appended to my testimony, I elaborate on the features required in assessments aimed at this purpose (such as detailed results, quickly available), so here let me simply elaborate on some of the implications of this view.

3. The governance structure for a new system of assessments, the National Education Standards and Assessments Council or NESAC, ought not to be so politically oriented as suggested in the Council report. Instead, if it is to be of help to teachers in improving instruction and learning, it ought draw on the expertise and organizational structures of of professional subject matter organizations, such as the National Council of Teachers of Mathematics, the National Science Teachers Association and the National Council of Teachers of English. This approach is sensible not just because the NCTM has been the source of the most widely cited example of subject matter curriculum and evaluation standards, but because if teachers are to be one of the prime users of the results of new assessments then their professional organizations ought to be

involved in the development of new assessments. Also, I might add that there appears to be widespread support for such a notion since 61% of a nation ample in 1989, when asked who should establish national achievement standards and goals, 61% said professional educators, while only 3% said that a consensus of state governors should set such standards and goals (Elam & Gallup, 1989, p. 44).

4. Moreover, if there is serious interest in evaluating the comparability of new assessments that might be used in different states and in judging whether new assessments meet "world class" standards then the NESAC ought to include people with expertise in test equating and in comparative education. Such members might be usefully drawn through nomination from groups such as the American Educational Research Association and National Council on Measurement in Education.

5. Also, rather than trying to invent more new kinds of assessments in ever more subject areas, I think that we ought to pay far more attention to assessments that already exist, even though they may not be directly keyed to national educational goals. What I have in mind here are assessments that engage students, and teachers (and ideally even parents) in learning activities spanning a variety of subject areas and that are pursued over a period of time. Examples of what I have in mind are the Westinghouse Science Talent Search and the Academic Decathlon.

6. Finally, rather than developing more new national assessments keyed to national standards,I believe that much more fruitful would be research and development on what is surely the most common means by which schools evaluate students and communicate student progress both to students themselves and to parents, namely school grades. Though grades can be and have been criticized on many technical grounds, it is worth noting that high school grade point averages have for several decades been shown to be better predictors of success in college than elaborately developed college admissions tests specifically aimed at predicting college success.

In sum, my reaction to the Council report is that unless and until the primary purpose for new national assessments is clarified, and until a clear vision of how assessments would work to improve teaching and learning is spelled out, then the future of such assessments will remain in the province of political rhetoric and will have little impact on teaching and learning in classrooms of America 2000. If new assessment are introduced, scores may well go up after a few years, but if we have learned anything from our national experience with educational testing in the last two decades, it is that increasing scores on tests and assessments may easily mislead us as to the underlying conditions of teaching and learning in our schools.

Haney Testimony, 2/4/92, p. 8.

References and recommended reading

Elam, S. 1990. The 22nd annual poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, September 1990, pp. 41-55.

Elam, S and Gallup, A. M 1989. The 21st annual poll of the public's attitudes toward the public schools. *Phi Delta Kappan*, September 1989, pp. 41-54.

Koretz, D. M., Linn, R. L., Dunbar, S. B., & Shepard, L. A. (1991). The effects of high stakes testing on achievement: Preliminary findings about generalization across tests. In (Ed.), Annual meeting of the American Educational Research Association Chicago, IL:

Madaus, G. F., & Kellaghan, T. (1991). *Student examination systems in the European Community: Lessons for the United States* (Contractor report). Washington, D.C.: Office of Technology Assessment, U.S. Congress.

National Council on Education Standards and Testing (1992). *Raising Standards for American Education.* A report to Congress, the Secretary of Education, the national Education Goals Panel and the American people. Washington, D.C. January 24 1992.

Schoenfeld, A. H. (1988). When good teaching leads to bad results: The disasters of "well-taught" mathematics courses. *Educational Psychologist, 23*(2), 145-166.

Chairman KILDEE. Thank you very much, Dr. Haney.

Mr. Goodling has another pressing appointment, so I'm going to have him begin the questions here.

Mr. GOODLING. I am going to allow all my time for all of these people who didn't have that magnificent privilege of serving on the Council.

[Laughter.]

I saw as my role serving on that Council as actually having two roles: to slow the process—and I think I did a pretty good job of that—and the equity issue. Those were my two hopeful contributions to all that we did.

I yield my time to everyone else who hasn't had that privilege and opportunity.

Chairman KILDEE. Thank you very much.

Mike, could you describe what the school delivery standards are in the report, and why the Council has recommended that they, unlike the other standards, not be developed on a national basis? We have national standards for the students but the standards for the schools we allow the States to develop on their own. Why was that done and why was that recommended?

Mr. SMITH. Let me start with the first part of your question, which was what are the standards. If you recall, in the House report and in the legislation, the Congress asked the Council to consider ways in which schools can be organized, to look at standards under which schools could be judged to determine whether or not they were providing equal opportunity and determine whether or not they were providing an opportunity to learn to students.

The Council responded to that charge from the Congress. The Congress, in its wisdom, saw that student standards were not enough. It implied in the legislation that we need to go beyond that; we need to go about thinking about how we can create an education system that provides opportunity to learn for all students. That means how do we set up a system such that all schools in the country can deliver to students the opportunity to learn the content standards.

That last sentence is very important, because this report implies something different than what's happened in the past. It implies that the schools would be held responsible, somehow—and it's not specified well in the report. It needs to be worked out. But the schools would be held responsible for delivering the opportunity to learn—the content standards, the standards that would be set out nationally—to all students. It would be triggered by the content standards. The content standards would be a goal. They would be a vision, high challenging standards for the Nation. How else do we imagine that we're going to move ahead unless we have got some sort of vision like that? That's the vision.

Now, how do you get there? Part of the way you may get there is to have the kind of national examination system that is laid out in the report. But that national examination system cannot exist in a valid way unless you have school delivery standards and the opportunity to learn in the schools that creates the opportunity for all of the students in the system.

Now, what could they be? Well, they could be two types of things. First of all, we can trigger a concern about the quality of a

school by looking at outcomes, by looking at the results of the examination, how well do students do on an examination that is specifically designed to measure the content standards. Just in exactly the same way that Jaime Escalante developed his strategy for teaching calculus, to teach to that examination. He taught to the examination, saying that exact same thing happens in every college calculus course in the country. You teach to an examination. There are examples of examinations in the libraries. You can go and study those examinations and you will know what you're going to learn. But then you have to learn it, and learning is the hard part. Earning takes time, particularly when it's a challenging content of this sort. So the assessment system would be geared to the content standards.

But, beyond that, there are other ways of thinking about what the delivery standards might be. For example, those students couldn't have learned calculus unless Jaime Escalante understood calculus and could teach calculus. There's a minimum standard for you. Why not say that all schools in this country should be populated by teachers who at least know the content and the content standards that we're laying out for students? We could even go one step further and even ask why not have those teachers be able to teach those content standards. I mean, that's a bigger step. It's not enough just to know the standards, to know the content standards, but you've got to be able to teach them. That's a fairly simple way of thinking about things, right?

Another way of thinking about it is to say: do the textbooks provide the kind of information that kids need to know to pass a test on those content standards? That's a very reasonable, very simple criterion. It's an input criterion. That has been criticized by lots of folks, who say we've got to get away from inputs and we need to go entirely to outputs. Let's go to simple inputs that have clear, definable relationships to the outputs we're talking about.

Teachers that are knowledgeable, that can teach the content and the content standards, textbooks that are of high quality and that reflect the content and the content standards, those are the kinds of things we're talking about, that I'm talking about, as school delivery standards. I'm sorry it took so long.

Let me just go to the second point, and that is the issue of why have national standards for students and somehow State standards for school delivery standards.

I don't see it that way. I see the national standards as being developed by national groups, organized, overseen by this entity we've been talking about—which is not the Federal Government. It's made up of people all across the States and localities, but not the Federal Government. We can't equate national and Federal in this case. That's what I see the student standards under the control of.

The school delivery standards, on the other hand, would be developed collectively by the States. Now, what is something that's developed collectively by the States? That certainly seems to be national to me. It's a different form of national. It is collectively by the States and then used by the States themselves to oversee, to audit the kinds of schools, to ensure the kinds of opportunities to learn that's called for in the report.

Chairman KILDEE. But we weren't very symmetrical—I served on the Council, too. We weren't very symmetrical in that. We said national standards for students, but then, particularly after the other end of Pennsylvania Avenue's influence came in, we changed midcourse, I recall. They were very opposed to any national standards for the delivery system.

You see, what I worry about is this. I can predict in certain schools what the test will be by how good those schools are. If I go into a school and find out that an English major is teaching trig, I can pretty well predict what those students are going to be doing in trig. Having been an English and Latin teacher myself, I could in no way teach trig. I could probably stay three pages ahead of the kids and that's about it. Or if there's a school system with 38 in the classroom and another one with 28 in the classroom, I can kind of predict.

I am just wondering why we set one criteria for student standards, and then really kind of switch gears in the middle of our considerations and say we won't do that, though, when it comes to schools.

Dr. Haney, do you want to comment on that?

Dr. HANEY. May I suggest an historical explanation?

Chairman KILDEE. Sure. I'm a history teacher also.

Dr. HANEY. Well, the history of testing in particular. Many times testing programs have set out to be used not only to make decisions about students but decisions about teachers and decisions about schools. Given the way the world works, it seems to me that students are always much less well prepared to protect their own interests than teachers or schools. I think this is simply a nice example of a large group—I'm suggesting that clear, high-quality standards are okay for the kids, but hey, we're not sure we want to meet them. I think the higher you get in the organizational structure, the easier it often is to avoid real meaningful accountability in terms of clear-cut objective standards.

Chairman KILDEE. Yet it was not the teacher organizations. I think the teacher organization wanted some standards for the schools there. It was another source that objected.

Dr. HANEY. Right. I'll bet I could guess who.

Mr. SMITH. I have one other comment.

I recall the question you asked me during the Council meetings, and that was wouldn't you prefer to have the school delivery standards developed at the national level. My response was yes. But I needed very desperately—I wanted very desperately to get the concept of school delivery standards into this report, deeply embedded in the way we begin to think about this process of changing American education.

Chairman KILDEE. Thank you, Mike. I recall our discussion of that.

We have such a great turnout of Members, I'm going to defer and call on, according to the custom of this committee, those Members in the order in which they appeared. Mr. Roemer.

Mr. ROEMER. Thank you, Mr. Chairman. I will be brief, to allow my colleagues time to ask questions, too.

I just have one question. We all talk about the high stakes. We're all back in Washington now after spending a month-and-a-half in

our districts, and we're back facing some of the same problems of the unreal world here in Washington. We get panelists on the one side that say here's a great idea, and the balance on the other side that say it's an awful idea, and then we go no where.

The high stakes here, ladies and gentlemen, are not only do 1400 SAT scores prohibit some people from going on, but look at the statistics of our schools right now, too—flat and declining SAT scores that are going no where; one-half of our students dropping out of school, or one-quarter of our students are dropping out of high school and another quarter are functionally illiterate when they do graduate from high school. One-half of the 18 year olds are failing to master English language skills, mathematical skills and analytical skills. All this adds up to about $3 billion a year in lost revenues. We had better do something, and as Members of Congress, my constituents want us to do something.

Let me talk about a program that's in Indiana. Whether this is called teaching to curriculum or a test or whatever, we have developed a program called "Indiana Plus." The business and the labor community got together at the local level and they went out, with students doing much of the analysis, and they asked future employers what are the three basic skills that you need from high school graduates to get jobs right upon graduation from high school. Three things came up almost on a universal basis: analytical ability, problem-solving ability, and computer skills. Even for labor jobs in the Midwest, it's not just hitting a hammer on a nail; it's being able to program some things on the assembly line.

What is wrong about our high schools being able to teach these three skills to our students so that they get good $15 or $16 an hour jobs on an assembly line, and they can keep our economy going and we can compete with other countries, and hopefully that is not being biased against students that have not performed very well up to a decent kind of pay.

Let me ask one on both sides here so I don't take up too much time.

Dr. HANEY. I would say that there's nothing wrong with that. I would say that a more appropriate national role would be, instead of trying to set up uniform standards, national standards, would be to promote the kind of role that is already played at the national diffusion network by organizations like the National Diffusion Network, where they critically examine innovative programs like that and critically examine evaluation evidence to see if the curriculum, the delivery and the assessments make sense. There's good evidence that that is a worthwhile program. And then information, not just about the assessment standards but the curriculum, the organization, and the delivery can be made available in other localities where they have precisely analogous needs. The point is to not disembody the assessments from the delivery.

Dr. RESNICK. I have to agree with what Mr. Haney just said, but I want to suggest that it's not in accord with what he said before, when he appeared to be arguing anyway that standards and assessment were going to be detrimental to this very process that Mr. Roemer just asked for.

There is an implicit analysis—true, it was not as well crafted a statement as perhaps the Council would have been able to do if it

had had a longer life—but there was an implicit analysis of why standards and assessment were desirable. It's an analysis that says the education system operates as a system, that it has to be coherent, and there need to be some clear targets for what everybody is aiming for and, in effect, some incentives.

Now, this doesn't mean high stakes of a simplistic kind, in which opportunities are cut off because you fail a test on a particular day. I want to set that straw person aside. It does mean what I think is a fundamentally new way of thinking about children's rights in this country to an education.

Right now we define children's rights to an education in terms of the right to attend school for a certain number of years, and for there to be certain funding that follows them. Implicit in—and what I would now like to put forward as the real point of the standards and assessment agenda—implicit in this argument is the notion that we can redefine what a child's right to education is. It is the right to achieve up to some agreed-upon level, not just to attend school.

Now, that is the kind of revolutionary idea that a lot of people are asking for. It has the power to bring about all of the other things that are being asked for, as long as resources also follow, both initially, probably in small amounts, and over time in much larger amounts. That's the argument. That's why it can work, because there will be, in effect, a new definition of the right to education.

Dr. HANEY. Could I just respond to that?

At a rhetorical level, that sounds very attractive. But I would simply point out that another widely recognized right for kids is the right to reasonable health care. Yet our Nation lags behind more than a dozen other countries in the world in the immunization of children. We know how to immunize kids, and we can do it cheaply.

So I ask, if the right to reasonable health care for kids has not led to the provision of a delivery that we know how to provide cheaply, why am I to believe that this defining of a right to education in terms of outcomes is going to make the delivery system follow.

Dr. RESNICK. We haven't established a right to health in this country. If we did, we would be behaving differently.

Mr. ROEMER. In conclusion, Mr. Chairman—and thank you for the time—I would like to say it shouldn't be an either/or situation, Dr. Haney. It is shameful when in this country we've seen a rising level of the number of mumps and measles cases in this country, when we are behind Ireland and Spain in infant mortality, 20th in terms of industrialized countries. That is shameful and we cannot look at this as an either/or situation. It should be both.

Thank you.

Chairman KILDEE. Ms. Molinari.

Ms. MOLINARI. Thank you, Mr. Chairman.

I'm a little confused, and perhaps I should direct this question towards you, Dean Smith. In your description of voluntary, when you talk about voluntarily opting into this program, would that be done on the State school, school districts

Mr. SMITH. Yes, State school, school district; it could be voluntary to any of them.

Ms. MOLINARI. So the States could opt in to national assessments and mandate that the school districts within the State apply?

Mr. SMITH. Exactly. Nobody is suggesting changing the way that the Constitution has set up the responsibility for education in this country. The States would have the responsibility. If the State doesn't opt in, doesn't make it mandatory to its local districts, a local district could adopt it or not adopt it.

Ms. MOLINARI. I suppose my question is, why would they? I mean, obviously there's a sharp penalty that would be brought about by a school district who opts for these standards, yet fails to reach these strategic assessments. Wouldn't it be safer to just not become a part of the program?

Mr. SMITH. I think that's right. There is a sense here, I believe—there was a sense in the Council—that if the standards are really well done, that it would be very hard for a State to not opt in, and it would be hard for districts not to come in in as strong a possible way as they could.

Let me give you an example. The National Council of Teachers of Mathematics, by themselves, working with mathematicians and with teachers and so on around the country over the last 8 years, have developed a set of standards. They are the first national standards developed in the United States. They're now being adopted State by State by State. Teacher training institutions are changing the way they're training their teachers; textbook publishers are changing the way they put together their textbooks, changing the content in the textbooks.

We are seeing a slow, quiet revolution in this country in mathematics education. It's because a group of people got together and decided, collectively, as a national group of people, that there is a better way to teach mathematics for our students.

Ms. MOLINARI. I appreciate that, and I don't disagree with your final point, that eventually States will feel the national pressure. But you also used the words there's a slow revolution, and I think we all agree that we can't afford to waste too much more time in terms of improving and revolutionizing our educational system.

Mr. SMITH. I agree. It may happen over time that we would get national standards in a whole series of content areas, but it would take 10, 15, or 25 years. I believe we've got to accelerate that process. The way to accelerate that process, I believe, is to get ourselves some sort of national body which can put on good quality controls. It can say, "Look, this was developed by the Nation as a whole," that these folks in geography went out and talked to people at the local level, talked to the people at the State level, and put together a curriculum that's really first rate, world class in some jargon. That's the kind of body we need to really accelerate this at this moment.

Ms. MOLINARI. Do you realistically think that that can be accomplished within the parameters of the political situation, a viable, vibrant governor not wanting to opt in at this particular point in time, being able to be challenged by this national council?

Mr. SMITH. I do. Yeah, I do. I think governors who are really interested in education will respond very positively to this. They will

argue very strongly that they believe the standards they have set up for their States are better than the national standards. And they may well win. There should be a competition among the States, in effect, to see which ones of those standards really meet the needs of the country. It's very important.

Ms. MOLINARI. Thank you.

Chairman KILDEE. Thank you, Ms. Molinari.

Mr. Sawyer.

Mr. SAWYER. Thank you, Mr. Chairman.

I hear everyone saying, in one way or another, that the real crux of all of this is what we measure is what we get, and we need to be very, very careful about what we choose to measure. In that sense, I hear the high stakes argument very, very clearly.

It seems to me that one of the things I have not heard—although I've heard the terms used, about the need for testing—the one thing I haven't heard is about the frequency of the use of these measurements, or whether they would be a turnstile event in a child's education at the end, or whether it would be used in all of it's multiplicity—I've heard that—as a tool not only to determine whether a child has met standards, but as a tool for diagnosis, remediation, or perhaps advance positioning of a child, if administered early and often in the child's education, before problems become so compounded over time by all of those self-fulfilling prophecies that they become unremediable. That's the first question.

[Laughter.]

Does the format anticipated in the testing and assessments that we're talking about today anticipate that use over time and that use that adds yet another variable to the potential uses for testing, that would reduce the high stakes rather than elevate them, and second, since what we measure is presumed to be what we get, and since we've talked a great deal about the importance of speed and how quickly we can act, has thought been given to the question of simultaneity in terms of the development of curriculum, expectations, standards, and the text instruments that are used to measure those, or is there a kind of sequence that is anticipated in the startup?

If what we're measuring, if the curriculum and the ability to teach—I just love that, the implicit right to learning and not just to attendance—if that is really at the core of what we're talking about, doesn't it imply that there needs to be at least at startup some measure of sequential development as opposed to simultaneous development?

Mr. Chairman, did I really take up all my time?

Chairman KILDEE. No, no. I thought you weren't noticing.

Dr. RESNICK. We have to have all of them in order to make good on this right to achievement. I think simultaneity is the right route to go. I am willing to sign symbolically—I didn't literally sign—a document like the Standards Council, one which mostly talks about the standards and assessment, because we can take them up one at a time, discuss them one at a time. I think if all we work on is the standards and assessment, we will get a punitive system and not one that is likely to raise achievement. So we've got to take it all on.

Those of us who are hardest at work on this, not waiting for legislation, are working on all of it, not just one the standards and assessment piece.

Dr. DARLING-HAMMOND. Could I add to that?

Mr. SAWYER. Sure. I would be happy to hear from all four of you at one time or another—but not simultaneously.

[Laughter.]

Dr. DARLING-HAMMOND. On your first question, the extent to which students will actually be helped by the assessment depends on two things, I believe. One is the extent to which we really do focus on the issues of how we can get schools to teach those things that Congressman Roemer raised—critical thinking, problem solving and so on. That does get back to the question of the delivery standards.

We need to be involved with solving the problem of a lack of qualified teachers in a lot of subject areas in this country. I mean, we have low math scores because a lot of math teachers aren't qualified to teach math. If we don't deal with that, having test scores isn't going to help us.

We need to deal with the fact that students in affluent majority schools have four times the chance of having certified teachers than students in low-income and minority schools. We need to deal with the fact that some schools spend $20,000 and some spend $3,000, and they buy different amounts of learning opportunities for that, including qualified teachers. So that's obviously going to depend on the one aspect.

The other is the extent to which the assessment is locally developed. If teachers are involved in developing with students portfolios, projects, a variety of assessment strategies, then those assessments are likely to be used. Again, of course, teachers need to be prepared to do these things well. Then the assessments are likely to be used in ways that will enrich what students, in fact, are offered in the way of learning, as well as what they learn.

To the extent that the further and further away it gets from the classroom the more far away the tests are in their development, the less likely it is that teachers will understand what it takes to improve students' chances of doing well on the assessments.

The question with respect to the high stakes is that if we use assessments inappropriately, we're likely to not get those benefits out of them. The question was raised earlier on how can we get people to adopt this kind of thing if they're likely to suffer penalties if they try. I think that's important. Dean Smith gave a good example with the National Council of Teachers and Mathematics Standards, of the power of a good idea to spread and to be used, to be adapted and brought into the system. That would not be happening if the NCTM standards had high stakes attached to them. It simply wouldn't be happening.

Dr. HANEY. My answer to your question gets back to what is the primary purpose of assessment. I think that we ought to be giving a lot more attention, if our aim is really to promote learning, to how assessments promote students learning directly—not because the teacher learned the results and taught them differently. I'm talking about assessments that engage kids and give them immediate feedback in a way that motivates them directly, ones that are

not necessarily keyed to academic subjects, but nevertheless can still comprise very challenging and high standards.

Let me mention two nice examples. One is the Westinghouse science talent search, which is dependent, incidentally, on a strong relationship between science teachers, knowledgeable science teachers and kids. The other is the academic decathlon, which engages teams of kids. Those, it seems to me, are nice examples of kinds of assessments which are not necessarily altogether locally developed but engage both kids and teachers and, ideally, even parents, in learning across subject areas and topics.

Mr. SMITH. I have a somewhat different answer than the others.

In my own view, a good teacher is testing all the time. They're asking a question and they're making an assessment about it; they're getting feedback and they're working with the child to put together a portfolio—whether it's called a portfolio or called their work for that week, or the work they're preparing to take home, whatever. I mean, all of those things are tests of some sort or other. They all give feedback, they all help develop the child, and they develop the teacher. They make the teacher more sensitive to the needs of that particular child and the needs of their classroom.

I think in the United States we ought to reduce the number of tests, the number of formal tests of the sort we're now talking about, to perhaps two during the entire time the student is in school, K–12. We overtest so much that we make the whole thing absurd. Kids laugh at it, they don't work hard on it, because the tests themselves aren't legitimate.

We have now national tests, standardized national tests that are given out to students all over the country, and they often have no relationship to the curriculum that the student is taking. Have you ever looked at an elementary school science test? It's absolutely bizarre. This is a random assortment of questions about science, facts about science. There is no in-depth attempt to get at an understanding of what the scientific process is, how science works in any situation or whatever.

The trick here, the thing being proposed here, is very different from the way the U.S. has gone about testing. The idea here is that these examinations, which would be an examination system, locally developed, State developed or whatever, they would be geared to the national content standards. They would reflect those content standards in just the way that the AP examinations reflect a syllabus that's created by a group of teachers and leaders. It's very different from anything——

Mr. SAWYER. Before we run out of time, can you talk just for a moment about the question of simultaneity or sequential development?

Mr. SMITH. I believe you've got to start with your content standards. I think you've got to know what you want the kids to learn before you can figure out how you're going to test it. You can have all sorts of tricks of testing. Portfolios and projects and so on, those are tricks. Those are ways of getting at the knowing. But you have got to know what you want the kids to know before you get at it.

The kids, in my view, in order for it to be legitimate, have to have the opportunity to know what's going to be on that test, in general, and to study for what's going to be on that test. It changes

the way we think about testing in this country. But the idea is, over a 4, 5, 6 or 7 year period of time, students will be preparing for sets of examinations.

Now, that wouldn't take most of their time in the class. They would be there just like students are normally now. But the teachers would be gearing it around the common content standards. When the kids got to the point in the fourth grade or eighth grade of taking this examination, they would be prepared because the teachers and the students knew all the way along what was going to be on those tests.

Mr. SAWYER. Thank you, Mr. Chairman.

[The prepared statement of Hon. Thomas C. Sawyer follows:]

STATEMENT OF HON. THOMAS C. SAWYER, A REPRESENTATIVE IN CONGRESS FROM THE STATE OF OHIO

Mr. Chairman, I want to thank you for holding these hearings this morning. I understand that you are planning two additional hearings to examine the Report and conclusions drawn by the *National Council on Education Standards and Testing*.

How Congress decides to act on these recommendations may well be the most important action taken in the educational arena this Congress.

Mr. Chairman, I must say that although I agree with many of the report's conclusions, I am greatly alarmed by what is missing. We have to be very careful here. If testing becomes one of the engines of educational reform in this decade we had better be prepared. Those of us who come from States where testing has already become a tool for making policy know that the issue is fraught with peril and consequences for individual students and communities.

What I wholeheartedly endorse is the development of national standards. This will take time, not a lot, but time. Then tests—as instruments—need to be very finely tuned. Only then, should we begin to *think* about using them on a national scale.

I have additional doubts, so I am looking forward to hearing the testimony of today's witnesses.

Thank you Mr. Chairman.

Chairman KILDEE. Thank you very much, Mr. Sawyer.

Mr. Gunderson.

Mr. GUNDERSON. Thank you very much, Mr. Chairman. I apologize to the panel for not being here during your testimony, but I have had a chance to review it.

One question that I would like some additional input on, if we are going to move to national standards—and I'm not saying I'm opposed to that—that makes the assumption that we have great disparity of State standards today. Particularly the two of you who were on the national council, can you give us some examples of the kind of disparity in State standards today that would mandate at least voluntary national—"mandate voluntary," those are interesting words——

Mr. SMITH. Suggest they're a good idea.

Mr. GUNDERSON. Yes, they would suggest voluntary national guidelines or standards would be helpful.

Dr. RESNICK. I can tell you something very specific, because last summer the group that I'm working with brought teachers from ten States, all of which have in place what's called performance assessments in writing. It's a simple idea. You're asked to write something for an hour or so, not "fill in the bubble" tests that are current now. So here we have ten States from all parts of the country who came together bringing with them actual examples of the

writing assessments their students had done, the actual papers the students had written. They came with a team of teachers who were experienced in grading the State exams. We worked out, with the help of an expert in these matters, Robert Wynn from the University of Colorado, a design in which teachers from California graded the papers of, say, South Carolina, according to the California standards. The South Carolina teachers graded the New York papers, and the New York teachers—you can get the picture. It needed a statistician to do it.

The results were astounding in two ways. The agreement was much higher than we had expected, given the fact that there had not been any effort to establish national standards before. But that agreement was on a correlation; that is, you could predict, in general, from how California graded a paper to how Illinois graded it, let's say. So in that sense, we appear to have in the country at the present time agreed-upon or shared standards for writing; that is, what the dimension in goodness in writing is.

But there was another surprise in the other direction. If you looked at the cutoff scores, what different States considered to be a passing level of writing, there was a difference of about 25 percent; that is, there are States that take a 25 percent lower cutoff point for what is good enough than other States. So in that one area, we appear to have agreement on the kind of thing that constitutes good writing, but not yet agreement on how good is good enough.

Now, these kinds of studies would have to be done subject matter by subject matter, or skill by skill. My prediction, although I give it with caution, since our predictions were wrong last summer, my prediction is that we would get less agreement at the present time on the dimensions of good math than we do on the dimensions of good writing, because there hasn't been as much time for this to happen.

But I have another prediction, and that is, by next summer, when these same States come back together, their cutoff points will have moved closer together, because part of the process of setting shared standards is the social process of examining each other's ideas and working them out together. That is very much going to be at the heart of this notion of comparability between different State assessments.

There is evidence from other countries, and now the beginning of it here, that you don't have to seek detailed statistical comparability in order to know that you've got comparable tests. This is going to be very much in our way, a negotiation process, one in which people of different opinions work with each other enough that they figure out what it is that they have in common. That's what I think this process will be like.

Mr. SMITH. Let me just make a comment.

There is great variability in the country. Wisconsin doesn't have content standards, curriculum framework, in the same sense that California has them, for example. It's much more up to the local district to decide what's important to them. There are some general specifications and there are general course requirements, but you don't find the kind of process that California went through to think about what kids should know in the area of mathematics and

social studies and so on in the State of Wisconsin, although Wisconsin is talking about it now.

Other States are also talking about it. South Carolina, for example, is moving to a set of curriculum frameworks. They are looking at what California did. They are looking at what New York does. They are looking at the NCTM standards. One of the things that's happening in the country is that people are looking at best practices. I believe that H.R. 3320, S. 2 and H.R. 4323, when and if they get passed, in those particular parts are going to accelerate that process. Not of just developing any old content frameworks. People aren't going to settle for that, because they're going to know that out there, there is something better. They're going to look to see what the best ones are and then they're going to tailor it a little bit to their own State.

What will happen eventually, I believe, is that we will get—I mean, this will take a long time if we don't have some sort of national body—we will get a growing convergence, as Lauren suggests, toward fairly high standards in the country. But we won't have a national body continually prodding States to upgrade the quality of their education, to upgrade the quality of their school delivery standards and so on. That's the advantage. We could get there, I believe, in the long run. It's a train that's rolling along already. You can get there in the long run just relying on the States. But I think we can accelerate it, accelerate it greatly, by putting into place some sort of national process.

Mr. GUNDERSON. Thank you.

Chairman KILDEE. Mr. Reed.

Mr. REED. Thank you, Mr. Chairman. I have not a question but a reaction, which I would like the panel to comment upon.

I will just advance the premise that if there's an appropriate test, however you define that, that implies that some children fail, and that in certain systems and certain schools lots of children will fail. It seems to me that some of the opposition to the idea of testing at any level—national, voluntary, mandatory, et cetera—is because people who make the decisions don't want to see children fail, because it reflects poorly on them, reflects poorly on the institution. But what has happened, really, by default, is that we have a system where they don't fail but few succeed, and those that succeed might do it despite the schools.

What I think is behind this political impetus to develop national testing is that that situation seems to be one we can't tolerate too much longer. People drop out of school and we find years later, when they enter the work force, they have no education or skills that are appropriate. That is why I think we're talking about national testing in some type of comprehensive way.

Now, I would just like your reaction to it, because I think we can engage in a pedagogic debate about what's the best tool, the content, et cetera, but the reality is are we going to get tough, frankly, and use assessments to evaluate, early on in the educational life of a child, whether they're making it, whether the schools are making it, whether the system is making it, or we continue to go forward and allow a system in which no one fails, no one succeeds, and we all get paid.

Dr. DARLING-HAMMOND. I think that to cast the problem as a lack of failure or explicit failure might misrepresent what's going on today. Many children do fail in school, and many of them are told they fail. They're retained in grade, they fail minimum competency tests, most States have them, and there are a variety of ways in which children's failure is repeatedly made explicit. And we do have a lot of high stakes testing as part of that certification of failure for students.

But that has not yet prodded localities, States, or the Federal Government, to ensure that the children who are failing have qualified teachers, have adequate resources, have textbooks, have programs that will ensure that they learn. I think the question in this dialogue, as well as any other, is are we going to do something differently with respect to the fact that children are not succeeding, in terms of holding governmental agencies, school districts, States, and the Federal Government, if it wants to get involved in this game, responsible for ensuring that students have the adequate opportunities to learn. We can certify student failure even better than we currently do, and the question really is, what are we going to do to hold ourselves accountable for delivering services to them.

I don't think that assessments themselves have to or, in fact, ultimately should be based on pass and fail to give good information about what students learn. I used the example earlier of the AP exams, where you get a score—one, two, three, four, five—or the SAT's, anywhere from 200 to 1400 or whatever. And there are certainly better ways of evaluating student progress than those which would still give information without having to make those judgments about students.

I would like to see the judgments made, and if we're going to have any kind of stakes in the systems applied to State governments, that don't provide equal opportunities to the students in their schools, or to school districts that allocate teachers inequitably to black and Hispanic students as opposed to white students in their districts, if we're talking about the stakes all being on the students and the local school, we're not getting at the source of unequal educational opportunity.

Dr. HANEY. I would just mention also that there's been a fair amount of research on not just the levels of dropouts but what kinds of programs are useful in keeping kids in school. Generally speaking, my reading of that research is the kind of programs that are productive in helping to keep kids in school and learning and engaged are not just some arbitrary barrier that says you've got to meet this; rather, it is some connections in the students' lives between their school work and adults, in school and out of school, who help them to see the social connections between their school work and their aspirations, not to pass a test but to do other things beyond school in life. It is not a test that's critical.

There I would go back to Mike's mention of "Stand and Deliver." It wasn't the test, per se, that was really crucial in that story. It was the teacher, not the test.

Mr. REED. I guess we're trying to get at this problem very indirectly because we have very little leverage on some other problems, like families that are not functional and other issues. We have to

recognize perhaps that this is the only lever we have to move the system. I think the sense is we want to try to move the system.

Mr. SMITH. I agree with that entirely.

Just one comment on Walt's final statement. It was both the teacher and the test, I believe. The test was there as a motivator for the teacher; it was a motivator for the student and it provided terrific rewards when the kids were successful.

That fed over. It didn't just influence the students that were in that movie, or that were depicted in that movie. It influenced generation after generation after them, because that particular school has retained a very high level of success in that area.

Mr. MARTINEZ. Would the gentleman yield on that point?

Mr. REED. Sure.

Mr. MARTINEZ. Do you know where Jaime Escalante is now? Is he still in that school, teaching that class?

Mr. SMITH. I don't believe he is, but——

Mr. MARTINEZ. I know he's not.

Let me ask you this. Have you followed up to see what has happened to that class and what has happened to the students there? I think you should, so that you could make a more accurate statement.

The success that was there, as this gentleman has pointed out, was Jaime Escalante, not the test. It was Jaime Escalante. In fact, the testers refused to accept the first test. They said the students cheated.

Mr. SMITH. I know that.

Mr. MARTINEZ. Why? It's that mind set, the mindset that you have. If you're looking to prove something, you can always prove it because that's what your mind is prepared to look for.

Mr. SMITH. I understand that, and perhaps I'm wrong. But my understand is also that the teacher next door to Jaime Escalante, a year later, had just as much success as he did, and it went on over a period of years.

Mr. MARTINEZ. For your information, you check and you'll find out that the high rate of success of Jaime Escalante has never been matched in subsequent classes, even in Jaime's classes.

Mr. SMITH. Let me pursue it one more step.

Isn't it true, though, that the school itself, over a period of years, had a high rate of success comparable to the rest of the country?

Mr. MARTINEZ. No. The problem in those schools—and let me explain to you why I know. It's because the school includes students from my district.

Mr. SMITH. I know it is.

Mr. MARTINEZ. Let me tell you what's happening. It's what this lady says. If you don't improve the school facility, you can set all the standards you want and nobody is going to attain them. The problem is that we still use, in most of our schools, as high technology, chalk and chalk boards. We're in a computer age. Those schools that have been fortunate enough because major corporations have donated computers, and because have been able to afford computers. They're moving ahead.

There was a case in California, Serrano, in which the direct relation of moneys expended for the students, which created a kind of facility for people to be able to learn, was the whole component of

that case, in which Serrano, the student, was found to be unjustly treated in his learning because of the inadequate amount of money spent in that school.

California has moved in a lot of ways to improve the school system. You were on the council that changed the whole structure of textbooks and what textbooks were supplied. California has been on the leading edge.

What concerns me is the fact that you've got a panel and a board and who's responsible to who? Most of them are appointed and not elected by the people. How do they interact with the authority of the local school board, who were elected by the local people and have the right to make a lot of decisions concerning this? There are a whole lot of unanswered questions in this issue that I don't think have been clearly thought out in what I've read so far.

I yield back to the gentleman.

[The prepared statement of Hon. Matthew G. Martinez follows:]

STATEMENT OF HON. MATTHEW G. MARTINEZ, A REPRESENTATIVE IN CONGRESS FROM THE STATE OF CALIFORNIA

We talk about standards—but we also need to talk about realistic opportunities to reach those standards.

The fact is that while technology is transforming the rest of our world, pencils and chalk remain the most advanced technology in many of our classrooms—and differences in school funding make it difficult for some of our schools to afford even that.

Experience in the private sector, studies by O.T.A. such as *Power-On* and results in some of our schools show that bringing technology to the classroom has the potential to strengthen our schools, to empower our teachers, to help even-out the resource differences among our schools.

If we are serious about improving educational standards *and outcomes* across this Nation we need to act now to support development of classroom technology and materials, to link and strengthen existing networks, and to develop effective materials and training. Classroom technology—firmly integrated with broader school reform—must be a fundamental part of education legislation. Thank you.

Mr. SMITH. I would welcome the chance to talk with you more about this, on both sets of issues.

May I respond to the questions of Mr. Reed for just a second?

Chairman KILDEE. Go ahead.

Mr. SMITH. The response I was going to give was that I believe it's a shared responsibility. In the same way that I've talked about school delivery standards in the past, Mr. Martinez, the notion that teachers know the content and know how to teach the content, you need the materials. It is also the responsibility of the student. It's a shared responsibility here. The student has to work hard in order to do well, no question about it. But they have to have the opportunity to learn the material.

If the material is well understood, it is a lot easier to begin to hold the institution accountable. So if we have a very clear set of content standards, we can then begin to hold the institution accountable about whether or not their teachers know the content and are able to teach it, and whether their textbooks—It goes right back to the Suranto decision. Are the materials there; does the school have the capacity to deliver; does the school have the capacity to deliver the opportunity to learn to all the kids? Those are tests that could be then made, once we have some idea about what we want the outcome to be.

Right now, in most of our States, we don't have any real idea what we want the outcome to be. It's sort of a general level on the SAT's and we don't really think hard about that, and the SAT really doesn't measure what we teach in our schools, or a general standardized achievement test. Again, it has almost no relationship to the curriculum in our schools.

This approach would have some relationship to that curriculum. You would be able to determine, at some level of confidence, whether or not the school could, in fact, deliver that opportunity to learn.

Mr. REED. Thank you.

Chairman KILDEE. Mr. Petri.

Mr. PETRI. Thank you, Mr. Chairman. I would like to thank the panel members for spending this morning with us and for the time they took in preparing their testimony.

I would like to associate myself with Mr. Reed's statement, and also explore a slightly different area. When we get into the problem of setting some sort of voluntary national testing standards, someone has got to come up with an appropriate test. I know there are groups working on that and they have come to some kind of consensus in the basics of reading, writing and arithmetic. But, even there, there will be different schools of thought as to the best way to teach these subjects, and, therefore, the best way to write these tests.

Do you think we're opening up what will be basically a contentious and not very productive area at the end of the day if we try to have national standards for civics, literature, history, government, and even geography? Where should we cut off voluntary national testing? Is there some core where there's a reasonable hope that responsible people of good will, from widely differing perspectives, can agree? Is there an area where it's unlikely we will get agreement; where we'll just get arguments and the lowest common denominator, or perhaps something that's counterproductive, which will set a lot of groups off in our society because they feel it's offensive? You know the problems we have with textbooks and everything else. We're going to get more problems with something called a national test if we're not careful.

Could you comment on that? Is that a problem? Is there a way of dealing with it at this stage of the game so we don't get into a thicket down the road?

Mr. SMITH. Let me respond quickly and then turn it back.

I believe you need to develop your content standards before you develop your test. You have to know what you want students to learn and you have to know what you want to teach before you begin to assess it. So it's the content standards that become the issue and that's where a lot of the controversy has been in the past.

I have two general comments. First, California is probably as diverse a State as there is in the union. It has gone through a process of developing science and social studies, history frameworks and literature frameworks, over the last 4 or 5 years. A contentious process but a very important process. In California, they ended up with what I believe are really first-rate frameworks. They're challenging, they're interesting. They embody sometimes controversial

theory, sometimes not so controversial. But they really put out a challenging menu for students to learn.

I don't believe that California went as deep as I believe we should go—that is, deep into the public, to have a really serious public debate about what should the content of our kids' schools be. It would be absolutely wonderful, in my view, if this country were arguing in a serious way about the content of the history that the kids should learn in schools. We don't do that. They do it in lots of other countries.

Mr. PETRI. But I sense a gap here, because people talk about national testing and the need for it. They talk about the problems of literacy, of people not being able to do basic things, and then, on the other hand, we seem to be wanting to get into all sorts of things that are really handled pretty well by college boards for that group of college-bound people who know roughly where they're going. Is it necessary for us to even get into that area? Can't we just stick to what we think the core problem is and then find various strategies to attack it, one of which might be national testing? I put that out as a premise, or as a question.

Mr. SMITH. I believe it's absolutely critical that we get into it. We can't wait until the student is in the 11th or 12th grade and taking the college boards. We need to be thinking about what the curriculum of elementary schools is.

We don't teach science in this country. No wonder we end up at the bottom of every ranking of science for 9 year olds and 13 year olds. No wonder. Our kids just don't learn it because we don't know what we want to teach. Many other countries know what they want to teach in science, so they can train their teachers to teach it. We don't know what we want to teach. Science is a big, buzzing confusion to an awful lot of teachers and to us, for that matter. I mean, we're no better. As a consequence, generation after generation comes out without knowing any science, without knowing enough science to really get into the courses en masse that they would need to do well in the college boards, for example.

What you see is a very small percentage of kids in this country taking college boards in math, in science, and very few advance placements. So we need to catch them earlier.

Mr. PETRI. Are there other perspectives on the panel?

Dr. HANEY. Yes. What's likely to happen, though this sort of widespread deliberation about what common standards are can be very useful from an educational point of view, there are two pitfalls that are likely to occur. First of all, there are going to be areas where there's not agreement. California is going to want to teach California history, and Vermont is going to want to teach something about the Green Mountains probably.

If there is great focus on only what is arrived at by a national consensus, that's going to mean that those things on which there is not consensus are going to tend to get pushed out of the curriculum. That's one problem.

The second is, that though there often is consensus on curriculum frameworks, such as the National Council of Teachers of Mathematics framework, please note that there is no national consensus yet on how to assess the learning of the mathematics curriculum laid out there. If you were going to have publicly-available

assessments, that's where we're really going to see some contentiousness about not just abstractly what people are to learn but how do you reasonably assess it, not just in the abstract but, as Dr. Resnick puts it, to use it in a thinking world. That's where the real problem is going to occur.

Dr. DARLING-HAMMOND. There will be contentiousness to the extent that we try to be very specific about what the right body of knowledge or set of skills apply in a particular way is. This is an area where the report, as you read it, takes kind of differing postures in differing sentences, as the extent to which we're really talking about a national curriculum, a set of national curriculum decisions with national assessments, or the extent to which we're talking about local and State curriculum decisions and assessment development within a relatively, one would have to assume, sketchy national framework for benchmarking in some sense. I think that is very slippery.

What is rarely understood, even in mathematics, is that the NCTM standards we have all been talking about present a different conception of mathematics than the AP calculus framework presents. It presents a different understanding of mathematics than the California framework, which itself is not the same as the New York State region's framework for higher level mathematics.

I think the question that people are going to have to struggle with, if they try to support a better assessment, which surely needs development, is how do we honor the various approaches to curricular standards, many of which all may be viewed as representing high standards in some sense, but different in some ways, as this kind of thing unfolds, to do that without dishonoring diversity, without dishonoring innovation and the capacity to move our educational system ahead.

Now, some other countries which have examination approaches of various kinds allow for a wide array of different approaches to the testing of mathematics. You can choose as a school to have your kids assessed with one kind of an assessment package or another, or to develop your own. There are certainly hints of that in the report, and I think how it plays out depends on which emphases one ends up relying on.

Chairman KILDEE. Mr. Owens.

Mr. OWENS. Mr. Chairman, coming last, I must apologize. I will be a little bit redundant.

First of all, I appreciate the level of discussion, the quality of the discussion. I want to congratulate you. I see that at least three of you, I think, are connected with federally-funded research centers——

Dr. DARLING-HAMMOND. I don't have one. Mine is not federally funded.

[Laughter.]

Dr. HANEY. Not me.

Mr. OWENS. Well, two anyway. That says we spend our money well.

[Laughter.]

Dr. HANEY. Notice which side the two are on.

Mr. OWENS. I hope we understand, though, in the process, that we owe it to ourselves and to our children impacted by this process

to be honest and to hold ourselves accountable. Some of the polite language here I think shows we don't understand how serious a matter we're into.

When you talk about delivery systems, you're really talking about a level of commitment of resources, a level of the education made being made at the State level or the local level or the Federal level. That's the problem.

Dr. Haney, you talked about teachers may not be in favor of delivery systems being assessed closely. No, the problem is not teachers. The problem is much worse. It starts at the White House. The problem is that the White House and the Secretary of Education and the Office of Management and Budget don't want to talk about assessing delivery systems because you get into the question of what resources do you need to make those delivery systems work. And when you get into the question of resources, it becomes clearer and clearer that States and localities will not have the resources to build world-class delivery systems or world-class schools. They're going to have to have resources from somewhere else. So we should not allow ourselves to be trapped into a process which ultimately may be phony or dishonest. We should force the issue of, as you build the accountability standards, you build in also a discussion and a statement of what resources are needed.

We have some situations right now where there are Federal mandates out there, Federal laws out there, such as the School Improvement Act. We are honored to have here with us the former Chairman of the Education and Labor Committee, Congressman Hawkins. The School Improvement Act is filled with a number of different mandates and requirements which would greatly improve the delivery system of schools, but they're being ignored. The Federal Government and the Executive branch has the responsibility to see to it that the law is carried out and that the schools are held accountable. Yet they ignore that because that might cost money, too. Those regulators that the President wants to get rid of cost money in this case. So the lack of commitment, the lack of resources, we have to discuss.

If you're going to establish world-class standards, the sequence, which one of my colleagues was referring to, the sequence has to begin with standards. Then you go forward and develop them as rapidly as possible. But it's immoral to not take the next step and assess the delivery systems or the education effort, the level of resource commitment, assess that before you move any further. Because to jump from the establishment of standards to the testing of students, which would create a lot of photo opportunities for a lot of ideologues and politicians but would not help the youngsters, it would be immoral. I think we have a moral responsibility not to move for the assessment of students before we have an assessment of the delivery systems, a way to assess what is it we're doing in order to create the maximum opportunity for youngsters to learn.

New York City's junior high schools, in all of the areas served by Latinos and blacks, does not have math and science teachers who majored in math and science. New York City doesn't have the capacity to cure that problem. You know, we can have a lot of math and science teachers as a result of the recession flowing to the schools—they know math and science but can't teach it; you need

real teachers—to train those teachers or to get those teachers in a time when we don't have a recession, the competition for math and science people is great. You need some Federal help. You need some dollars. We have to discuss those dollars.

My question is—I've already answered it.

[Laughter.]

Would it not be immoral or unethical——

[Laughter.]

Would you care to comment on my statement?

[Laughter.]

Mr. SMITH. It's easier to answer your question, and I think everybody here would answer "yes."

Mr. OWENS. Maybe the other part is more relevant.

Certain people have said—I think Governor Roemer has been quoted as saying even if Congress does not create the standards and testing panel, we should go right ahead, that the goals panel would do it anyway.

Would you comment on the movement ahead of that process, leaving out the Congress, which represents all the people in ways in which the Executive branch doesn't, how likely we are to succeed if we don't have that kind of consensus at the highest levels?

Mr. SMITH. Let me make two comments. First, I agree entirely with what you said about the need for resources. I think Linda's phrase was it would be a cruel hoax. There's a phrase in the report which talks about a moral imperative to provide educational opportunity for all children. We had many discussions of this in the Council, as the Chairman will attest. Many of us come down very strongly on your side, Mr. Owens. I believe you're absolutely right.

That's what underscored the notion of school delivery standards. We needed to get that concept into the document so that we could begin to use it as leverage.

Mr. OWENS. I congratulate you for getting it in. I think it was probably very, very difficult.

Mr. SMITH. But we succeeded.

Now I've forgotten your second—Oh, should the Congress go ahead. I believe it's absolutely critical that the Congress is part of this effort, absolutely critical.

Mr. OWENS. Should the goals panel and Executive branch go ahead with the encouragement of Congress?

Mr. SMITH. Let me go to the opposite side of that. I think it's a mistake if it's only the goals panel. I think we need the efforts of the Nation, and the Congress is absolutely critical when one thinks about the Nation. The Congress, the administration, the governors, the chief State school officers, people who represent us collectively across this Nation, school board members, all those people have to be involved at some level. And they could be.

My view on how you get that is for the Congress to create some sort of independent body that can operate and reflect the views of the Nation, that can hold people accountable. I think it's really important.

Dr. HANEY. First, on your statement, I entirely agree, and Mike seems to agree, and I think Lauren, in essence, said she agrees. If everyone agrees, then it seems to me there is an obvious recommendation that the panelists all could make to the committee;

namely, that the committee not pass legislation endorsing the Council recommendations without specifically suggesting that a priority ought to be the development of the school delivery standards rather than student standards. It seems to me an obvious recommendation of the panel unanimously, given the comments I've heard.

More generally, however, regarding your second question, should the Congress be involved or, if not Congress, then who, it seems to me that right now there is a terribly important need for national leadership in the realm of education. I would hope that more leadership comes forth from Congress on educational issues.

If you look at what citizens say about who should be making decisions about standards in education, about 3 percent nationally say it should be the governors, or a consensus of the governors, and a similar percentage says the Federal Government. Most citizens want local government bodies and professional educators to be the ones making the concrete decisions about national standards, not the Congress. So the support is yes, for some structures. But it seems to me that you've got to maintain the involvement of professional education groups such as the National Council of Teachers of Science, the National Science Teachers Association, the National Council of Teachers of English, rather than setting up totally separate bodies that are unconnected with ongoing professional organizations. Then on the public input, groups like the National Parent Teachers Association, the various advocacy groups.

It seems to me it's important not to try to set up altogether new structures but to draw on the structures that are already out there involved as education advocates in the States and communities.

Dr. RESNICK. Let me respond to your implied question, as well as the direct one.

I think Congress needs to act and be very much part of what we move ahead with. I think school delivery standards are crucial, but we need the whole package of this atlas, of the kinds of standards that were described in the Council report. That is because of this notion of the right to an education up to a standard of achievement.

School delivery standards without that built in will not be enforceable. They will not have any meaning. They will be squabbled over endlessly and there will be no way of finding out whether we have set the right school delivery standards. So it's the whole package that we're going to need, and it's this right of our children to an education that really is the clue to why we need it all.

Mr. OWENS. Thank you very much.

Mr. Chairman, I would like to note that I would like extra copies of this section of the transcript.

[Laughter.]

Chairman KILDEE. Thank you very much.

I'm going to go back to Mr. Petri for a question. He wants another round of questions.

Mr. PETRI. I agree. It was a good speech.

[Laughter.]

Actually, I have a question that I didn't get a chance to ask during my 5 minutes. It focuses in a little different way on what my colleague was getting at. If we do have national voluntary test-

ing for students, and we get results which show that some do well and some don't, how far does that get us if we don't have any measures to indicate why that is? That is, it could be the teachers, it could be the administrator of the school, it could be the parents, or it could be the inherent difference of ability that we all have, whether we're children or adults.

If we have a test and the kids are doing poorly, then what? Maybe the teacher doesn't know anything and can't teach. But the tests won't tell us that because there could be other variables. Maybe it's difficult to teach a certain class. If we just focus on the performance of the kids and blame the teachers, we will discourage teachers from doing what you wanted them to do, which is to teach hard-to-teach communities. Therefore, it's not fair to look only at one variable in isolation, because it doesn't tell you, in and of itself, what to do to correct the poor performance.

Could you comment on that? In other words, should we be looking at national voluntary standards for teachers and for principals and for parents as well as for students?

Dr. DARLING-HAMMOND. I think the utility of the kinds of assessments that the report is trying to aim at is that they be embedded in a local system of continual school evaluation and school improvement such that they, in fact, trigger inquiry into what is going on with a particular student, what's going on with a particular school, a set of classrooms and so on.

Now, that kind of process can't happen just mechanistically, as you suggest. You can't just say well, here's the test scores for this school, and if they're not high enough, we're just going to do something to punish the teachers or take money away from the school or whatever.

I am reminded when I think of that of a wonderful school in East Harlem, which is achieving a 100 percent graduation rate and sending most of its kids on to college, and 99 percent of the kids are local students from this predominantly black and Latino community. A third of them are mainstream from special ed. Now, assume that the director of that school would oppose some of the proposals around consequences based on national assessments, despite the fact that they've already put in place the kinds of assessment strategies called for here, because it would prevent her from being able to accept all of those special education students every year. If she had to guarantee a greater and greater proportion would always meet a higher and higher standard, she couldn't offer the high quality of education to the students who really most need it. So you can't use a mechanistic strategy for penalizing those educators who are working hard to overcome the other kinds of challenges that students may face in a variety of ways.

What you want to have happen is for the assessments to be part of a school-wide inquiry process. In the statement that I entered in the record earlier, for purposes of time I skipped over one part which said that what we want is a system whereby schools will engage in regular kinds of inspections or audits and presentation of themselves in what they do to the community so that you can use it in a diagnostic way. You can develop other procedures—teacher evaluation systems, school-community dialogues of various kinds,

improvements in school delivery standards and so on that have to be there. But it can't just be a mechanistic use of the test scores.

Mr. SMITH. I agree with what Linda says, but I want to make one comment on it.

This approach of a school that's responding in a creative way to information, information from the test, information from other sources, is particularly useful when there are clear ideas about what you want the school to accomplish, which brings you back to the content standards idea.

Let me also use this opportunity to say one other thing. I was one of the two people on the Council that signed the statement that Linda read to us. I did it because I agree with it. I agree with it. I believe there's an absolutely remarkable congruence between what that statement says and what the intent of the Council was, and what much of the Council report says. And let me just run through very quickly a set of points.

The statement says there shouldn't be a single national test. The Council came out explicitly against a single national test. The statement says there should be challenging educational standards and even implies that they should be national. Clearly, that's what the Council report says. The statement says there should be strong community input in developing standards. Absolutely. It goes right along with the Council.

The statement says there should be a strong reform program in support of standards and assessment. Clearly, that's what the Council report says when it's talking about systemic reform. The statement says no inappropriate use of tests for accountability. Who could disagree with that? It is very clear in the Council report that we should have legitimate accountability, that there should be the opportunity to learn, and then you use a test for accountability along with other instruments.

Finally, the statement implies that there should be something like school delivery standards—although they don't use the terminology. The points of congruence are extraordinary for two documents that were developed totally independently, as far as I know.

Dr. HANEY. Mike, I would just say that that's your interpretation of the main points in the Council report. There's a lot of ambiguity and a lot of contradictory language. For example, in the Assessment Task Force report, which says you can't use one test for all these different purposes, and then you read the text of the Council report and it says it "may be" that one test can't do all those things. There are a lot of different messages that come through.

Dr. RESNICK. But we're not trying to do an exegesis on a test. We're trying to think about what the right policy is for the next step and the policy gets worked out over time. What that document did was establish a platform from which Congress now, and if we move ahead with either what the Standards Council recommended there or some alternative body, can deal with all of these kinds of questions.

One cannot expect a group that had 6 months to work to figure out the details of education policy over the next decade. What can be done is to set some broad principles out. It is exactly this kind of interpretation that is needed to use a document like that. So this is precisely the right kind of process.

I want to add that not everybody, because of travel schedules, got a chance to sign that document in the first three or four days. There is at least one more, myself, who was a member of the Council, who has now signed it. There might be others.

Dr. HANEY. With regard to the broad principles, though, I guess there's a fundamental disagreement, it seems to me, in what is laid out in the Council report about the relative priority for school delivery standards versus student assessment standards. What I heard today from every member of the panel is that school delivery standards to prevent a cruel hoax in assessing students need equal or maybe greater priority than student assessment standards.

Dr. RESNICK. Equal, not greater, because the whole program cannot work if there is not real agreement on what it is we're after.

Dr. HANEY. That is a fundamentally different message than comes through in the Council report.

Dr. RESNICK. So that's the message we're giving today.

[Laughter.]

Dr. HANEY. I just wanted to clarify that it was a somewhat different message than the Council recommended.

Chairman KILDEE. As I mentioned earlier, I didn't think it was quite a parity either on the school delivery standards as it was on student standards. The two members who were on the Council, we were in kind of a cabal there trying to strengthen that part of the report.

Let me ask you this question. What role should the Federal Government play—and this is getting off into the future a bit—but what role should the Federal Government play in helping schools reach these school delivery standards, however they are determined by a group of States? What should the Federal Government, once those standards are set, assuming they are set, what should be our role? Dr. Darling-Hammond?

Dr. DARLING-HAMMOND. One major role that I think the Federal Government ought to play in terms of school delivery standards—we have said over and over again that teachers and what they know is a major part of the key—is to establish and reestablish the kinds of programs to recruit and highly train and improve the education of teachers across this Nation. We lost the Urban Teacher Corps in 1980 and we have serious imbalances in the capability of teachers across this country to do the job. That's a role that the Federal Government has played with respect to delivery standards, if you will, or capacities of schools in the past. It's one that the Federal Government has played and continues to play with respect to the recruitment of an adequate supply of well-trained physicians in this country, and it certainly is a foundation for any of the rest of this to begin to happen.

There is certainly as well opportunities for the Federal Government to be much more proactive and clear about what other components of a quality education are and investments to help secure those in places where they're not currently secured. Because the major mission of the Federal Government has traditionally been, at least since 1965, to equalize the opportunities available to students across the country.

As we're talking about the importance of high universal standards to national competitiveness and national well-being, that is a fundamental component of the national interest that the Federal Government ought to consider assuming or reassuming.

Chairman KILDEE. Mike?

Mr. SMITH. I think you've got a good start in H.R. 3320. What is really needed is a coordinated effort between the Federal Government and the State governments. They and they alone have the capital, have the resources to really provide it down at the local level, so that the local level has the energy, enthusiasm and the structure to really make a dent in our problems.

H.R. 3320 gives you that opportunity. It gives you the chance to say to States, what we expect you to do is to go out and carry out some systemic reform, begin to tie together some various elements that are really critical, around the kinds of teacher training and curriculum materials that are necessary for these kids to learn.

Along with H.R. 3320, of course, you need to modify title V of the Higher Education Act, which speaks to the issue that Linda spoke to. There is a terrific amount of money from the Federal Government that goes into teacher training, but it's not coordinated in any serious way. The money in the Department of Education, the NSF, money in a variety of other places, has to be coordinated through something like the H.R. 3320 mechanism.

Chairman KILDEE. Dr. Haney.

Dr. HANEY. Elaborating on Linda's comments and Mike's comments, I would point out that it seems to me, historically, and indeed currently, the Federal role in helping to ensure equal educational opportunity probably ought not to be viewed as strictly a cooperative one, because, unfortunately, there are some places where at the State level we do not see State level leadership promoting equal educational opportunities.

Though I think it's regrettable, I think there still is an important role for the Federal Government to play in providing concrete incentives to make sure that States do provide minimal levels of equal educational opportunity.

I agree with the emphasis on doing more in teacher education. The other thing I would mention is that mechanisms ought to be put in place to ensure that any initiatives supported by the Federal Government with regard to the National Council recommendations ought to be closely coordinated with other kinds of assessment activities funded by the Federal Government—for example, in the National Center for Educational Statistics, the National Assessment of Educational Progress. Not that those data-gathering assessments should necessarily be totally comparable with any new assessments that come out of this initiative, but that, at a minimum, we have well-documented evidence as to what's covered where and what's not covered where.

Chairman KILDEE. Dr. Resnick.

Dr. RESNICK. I agree that H.R. 3320 is a very good start. It lays out the principles of a systemic reform agenda. Whether direct further entitlement funding from the Federal Government will be required to have the effect that H.R. 3320 is seeking is going to have to be addressed over the next period of time.

I want to also add that there's a classic role that does need to come from the center, and that is the research base for all of this. We're talking about a very new kind of educational process. There are going to be mistakes along the way. We need to have a learning system in the reform process itself.

I know that OERI is coming up for reauthorization. That ought to be considered as part of this total. There are recommendations as much as 1 percent. That doesn't sound very high, but it's a lot higher than what we are now spending of Federal money on research and development. No field that is trying to move ahead quickly—take any industrial company that's in the lead as an example—is going to be spending less than 1 percent on studying itself and improving itself along the way.

Chairman KILDEE. Dr. Haney.

Dr. HANEY. Forgive me. I have to catch a plane, and I have to teach a class this afternoon.

Chairman KILDEE. I understand.

Dr. HANEY. It's been an honor and a privilege.

Chairman KILDEE. Being a former teacher, I understand that very much. Thank you very much.

First of all, I want to thank all the witnesses for joining us here today. It is particularly good to see Mike and Lauren again. We worked in the "vineyard" on the report. There wasn't always unanimity there, but it was a very good operation, all in all. We will look at the report and see what legislative addenda is needed or things we might want to do to it.

Your testimony today has been very excellent. It will give us a chance to carefully review the report of the Council. We've created a good record here. I would like to keep the record open for probably 2 additional weeks for any additional testimony that any of the witnesses wish to submit, or that other people wish to submit. Mr. Martinez also asked permission to submit a question to the panel in writing, and we will get that question to you.

Unless there is something further, we will stand adjourned. Thank you very much.

Mr. SMITH. Thank you, Mr. Chairman.

Ms. RESNICK. Thank you.

[Whereupon, at 12:04 p.m., the subcommittee adjourned.]

[Additional material submitted for the record follows:]

Donald M. Stewart

Statement To House Committee On Education And Labor, Subcommittee On Elementary, Secondary And Vocational Education

The College Board appreciates the opportunity to comment upon the issues raised in the report of the National Council on Education Standards and Testing, and particularly on the desirability and feasibility of national education standards and testing. We applaud the efforts of the subcommittee in bringing these issues forward for Congressional discussion.

As a national membership association of more than 2,800 schools and agencies in higher and secondary education, the College Board is in a unique position to offer these comments. Throughout our 90 year history we have been committed to improving academic standards through collaborative efforts involving school and college faculty. The Board sponsors in guidance, admissions, placement, assessment, financial aid, and credit by examination programs to assist the school-to-college transition of some four million students each year. The Board also sponsors programs to help schools prepare all students for college or the world of work.

This is a time of extraordinary opportunity for education thanks to the leadership and vision of many policymakers, including the members of the House Committee on Education and Labor, and especially its chair, Representative William Ford, the chair of this panel, Representative Dale Kildee, and the ranking member of both the full and subcommittee, Representative William Goodling. For the first time in our history we have a set of national education goals. Achieving these goals can improve the quality of American education not just for some, but for all. And achieving these goals is imperative since the stakes are high and the end results will determine our educational future and even our status as a democratic nation. It is in this context that national standards, and a national assessment system built on these standards, are both desirable and feasible.

Success in reaching national education goals will occur only if:

- **there is a consensus on standards and what students are expected to learn -- the "what" of schooling;**
- **teachers and schools are empowered to support all students in achieving these high standards -- the "how" of schooling; and**
- **there are appropriate means to measure, both in a formative and summative manner, educational progress -- the "how well" of schooling.**

Furthermore, we must take steps along the way to ensure that all students are expected to meet high standards and have access to quality education. High standards without equity are neither productive nor fair. Instead they are destructive and morally unfair.

The College Board is doing a great deal to advance this vision and achieve the national education goals.

BEST COPY AVAILABLE

ESTABLISHING STANDARDS, THE "WHAT" OF SCHOOLING

The distinction between curriculum -- the "what" of schooling -- and assessment -- the "how well" of schooling -- cannot be blurred. Agreement upon the "what" of schooling must precede agreement upon ways to measure the "how well" of schooling.

In that regard, the College Board is pleased that both the National Education Goals Panel (NEGP) and America 2000 acknowledge the importance of first establishing standards. The NEGP endorses the creation of a framework reflecting "what the nation wants young people to know and be able to do as a result of their years in school" as the first sequential step in the development of their proposed examination system. The America 2000 strategy calls for the development of "new world standards" and a new nationwide examination system tied to these standards.

The College Board endorses the underlying assumptions and recommendations of the report of the National Council on Education Standards and Testing, Raising Standards for American Education -- that standards are desirable and feasible; that they be voluntary, not mandatory; that they be federal, not national; that they involve broad-based participation; and that they be dynamic and continually updated.

The College Board is well-positioned to assist in the standards setting process as outlined by the National Council on Education Standards and Testing. As an association of secondary and higher education institutions, the College Board has a wealth of experience in bringing together educators from all levels to reach agreement on educational objectives and test content across a wide variety of academic subjects. Our standards setting activities resulted in a publication -- Academic Preparation for College (also known as the "green book") -- describing "what students should know and be able to do" upon graduating from high school. The "green book" describes learning outcomes in English, the arts, mathematics, science, social studies and foreign languages. It also identifies and describes basic academic competencies -- reading, writing, speaking, and listening. A subsequent set of College Board publications, the "rainbow" series, provides specific curriculum and instructional suggestions about how to achieve the results outlined in the "green book." In fact, the math community was greatly assisted in its consensus-building and standards setting by the College Board's publication of Academic Preparation for Mathematics.

These efforts at standards setting are powerful because educators and educational associations are involved. This is as it should be. The College Board hopes that future standards setting activities will build upon, and be informed by, these past efforts.

EMPOWERING TEACHERS, THE "HOW" OF SCHOOLING

Standards are translated into learning through the teaching process. The current discussions of national assessment have paid too little attention to the critical role of teachers and other

educators in implementing standards and there can be no success without this involvement. The assumption is that teachers should and will respond to the national education goals; however, more must be done to include them directly in the process and to support them in their classroom efforts.

New standards will require teachers to improve current skills and develop new ones, to invent new methods and to experiment. This will not happen accidentally or indirectly. It must be planned and executed. Professional development and training efforts must be put in place to involve teachers in the setting of standards and to assist them in preparing students with a full range of abilities to meet outcome expectations.

CHOOSING ASSESSMENT INSTRUMENTS, THE "HOW WELL" OF SCHOOLING

The education process described above starts with the establishment of standards, involves teachers in the translation of these standards into classroom and other learning activities, and culminates with the assessment of learning outcomes.

This final phase of the process leads to a number of questions. Should there be a national examination (or system of examinations)? What kinds of testing should be utilized and developed? What are the purposes of testing -- accountability or individual learning?

The College Board is addressing these questions and taking steps to answer the current call for new assessments. Last June the Trustees of the College Board authorized the development of a major initiative -- an array of secondary school syllabi and related assessments, supported by professional development opportunities for teachers. We call this initiative Pacesetter, signifying our hope that they will indeed set a faster pace for educational reform than has been the case up to now.

Pacesetter is based upon the following premises:

- The nation can only meet higher educational goals if teachers and subject matter experts are active participants in the effort.
- True reform must start with precise definitions of the specific skills, knowledge and understandings that all students should possess to compete successfully in the American workplace.
- Those definitions must be in place before educators decide how to measure progress.
- Teachers must be prepared to successfully transmit the skills, knowledge and understanding reflected in the definitions.

- Assessment should allow students to demonstrate their achievement in a variety of ways, including performance-based tasks.

- For student development purposes, assessment should be embedded in instruction.

Based on these premises, Pacesetter will provide states, school districts, and schools with a comprehensive, integrated and professional approach for setting higher academic standards, improving instruction, and assessing student achievement. Pacesetter assessments will be developed after new standards in each subject and their related curricula and instructional practices are in place. Both evaluative classroom activities and end-of-course achievement tests will be incorporated into Pacesetter.

Currently the College Board is working with the Mathematical Association of America (MAA) and the National Council of Teachers of Mathematics (NCTM) to develop a mathematical experience for the upper high school grades which will be designed to implement the NCTM standards in school and classroom practice. The NCTM, the MAA and the College Board have agreed to work together towards a mutually acceptable program of mathematics syllabi, teacher development and assessment. The initial Pacesetter offering in mathematics is projected for 1993.

Similar agreements have been recently concluded with the National Science Teachers Association, the National Council of Teachers of English and the American Council of Learned Societies (for world history). Together the College Board and these disciplinary organizations have created nationally representative task forces of teachers for each subject area. The end result will be the identification of essential learning outcomes in the curriculum, and ways to attain them, and the development of a process for assessing student achievement that integrates the daily work of students with the end-of-course examinations.

An important College Board antecedent for Pacesetter is the Advanced Placement (AP) program, described by Secretary Alexander as "the best test we have today to honor proficiency in major skills." In fact, President Bush is proposing to use performance on AP examinations as the basis for the Presidential Citations for Educational Excellence, and AP data is also used as an indicator on the 1991 Report Card issued by the National Education Goals Panel.

The current AP program provides students with the opportunity to complete college-level studies while in secondary school. It is curriculum-based, with teachers directly involved in the development of courses and the development and grading of examinations designed to measure achievement utilizing both multiple-choice and student-produced answers as forms of assessment. Multiple-choice questions are included to sample broadly from the content of the course, while the free-response format enables a student to demonstrate the ability to "do" some valued aspect of the subject.

However, it should be noted that Pacesetter addresses learning at the high school level for all students, while AP addresses college level study for the most able high school students.

Finally, the College Board is pleased that the report of the National Council on Education Standards and Testing clarifies the purposes of testing and distinguishes between testing for accountability and testing for individual development (learning). In practice there is frequently a difference between the two, and this must be understood by all those concerned with improving education.

Testing for accountability involves the gathering of information about the performance of groups of students in order to inform educational policymakers -- including government officials -- about the effectiveness of schools at the state, city, district and school level. Testing for individual development provides information to teachers and students that will help guide instruction and learning in the classroom; it provides feedback to students, parents and teachers so that instruction can be shaped to help overcome weaknesses in learning.

Tests for individual development differ from tests utilized for accountability in two important ways. A test used for accountability must cover a broad range of content representing a good cross section of valued educational objectives. In contrast, a test used for individual development is narrower in scope but deeper in content; and a test must provide information to both the teacher and the students about the mastery of very specific subject matter by individual learners. Furthermore, tests for learning must be given at the individual student level. This is not true for tests of accountability; carefully selected samples can provide sufficiently reliable data for this purpose.

Thus, the College Board agrees with the recommendations of the National Council on Education Standards and Testing for a system characterized by multiple assessments, rather than a single test. We agree that such a system should be voluntary, evolving, coherent, consistent with national standards; the tests should be technically valid, reliable and fair and utilize new assessment techniques. We are pleased that Raising Standards for American Education acknowledges these characteristics as the basis for a new system of national assessment.

ENSURING HIGH STANDARDS FOR ALL STUDENTS

The College Board wishes to underscore the need to guarantee access and equity for all students. At a time in our history when cultural, racial and ethnic diversity has never been greater, it is essential as a matter of both social justice and economics to assure quality education standards for everyone. We must not allow the rush towards the establishment of standards and the creation of new assessment systems to obscure (or worsen) the extreme inequalities of opportunity that currently exist for minority and disadvantaged students. We must monitor every step, every plan,
and every action so that equity is advanced while high standards are achieved.

It is appropriate to mention in this context a national campaign launched by the College Board, Equity 2000. Designed to raise the college-going rates of minority and urban students by the year 2000, the program is based on research that statistically demonstrates the importance of student mastery of algebra and geometry as a prerequisite to pursuing and succeeding in higher education. The setting of standards, the preparation of teachers to meet those standards, and the utilization of assessment techniques to identify strengths and weaknesses in student progress are all an integral part of our Equity 2000 project.

We cannot allow comparisons with the educational achievement of other countries, and their testing programs, to obscure the real strength of American education. We must remember no other country has an educational establishment which aspires to the degree of student participation, the pluralism of means, or the multiple layers of opportunity which exist in the United States. The achievements of other systems come from the fact that they have identified clear educational objectives toward which instruction for some students is aimed. The challenge for America is to establish objectives which are no less rigorous, but also to create the means by which they can be pursued successfully by all students.

CONCLUSION

The College Board very much appreciates the opportunity to submit comments to the Subcommittee on Elementary, Secondary and Vocational Education. It is our hope that our comments on the importance of setting standards, the preparation of teachers to help students meet those standards, the use of specific assessment instruments and the importance of ensuring equity will be helpful in your legislative considerations.

It is the fervent hope of the College Board that the commitment now being focused on educational reform will bring positive results for students, their parents and schools. Through dialogue, consensus and cooperation American education can surely emerge strengthened and revitalized.

p:\wplib\housecom.iks
March 4, 1992

OVERSIGHT HEARING ON THE REPORT OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING

WEDNESDAY, FEBRUARY 19, 1992

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON ELEMENTARY,
SECONDARY, AND VOCATIONAL EDUCATION,
COMMITTEE ON EDUCATION AND LABOR,
Washington, DC.

The committee met, pursuant to call, at 9 a.m., Room 2175, Rayburn House Office Building, Hon. Dale E. Kildee, Chairman, presiding.

Members present: Representatives Kildee, Martinez, Lowey, Unsoeld, Mink, Pastor, Goodling, Klug, Petri, Gunderson, and Henry.

Staff present: Susan Wilhelm, staff director; Jeff McFarland, legislative counsel; Damian Thorman, legislative assistant; Lynn Selmser, professional staff member; and Andy Hartman, education coordinator.

Chairman KILDEE. The subcommittee meets this morning to continue a series of oversight hearings on national standards in testing. These hearings are structured to promote an extensive discussion of the various issues related to the subject of national standards in testing. The views and perspectives of the witnesses will enable the subcommittee to more fully understand the most critical issues and respond more effectively.

I'm especially pleased this morning to be releasing, along with Mr. Goodling, a report from the Office of Technology Assessment called "Testing in American Schools: Asking the Right Questions." In response to a request from the committee of the Office of Technology Assessment, OTA has conducted a 2-year extensive review of educational assessment in the United States and abroad.

The report will provide valuable input into the committee's deliberations on the important issues related to the establishment of national standards in testing. The Senate, as you know, has passed an education bill this year and has a section on national standards in testing. This committee is laboring over a number of bills, including how we will deal with the report of the Council on that. Mr. Goodling and I served on that council.

At this time, I'd like to recognize my good friend, Mr. Goodling, for an opening statement.

Mr. GOODLING. I don't have an opening statement. She said yes I do. It's called just-in-time delivery; right? That one I read at midnight, and I crossed out one line. I don't know whether they put it

back in or not. I'll just submit that statement for the record. This morning I was handed a news article saying Montgomery County reading and math scores rose.

Of course, included in the article is what I've been saying from day one, that we need to test to find out where we failed as a teacher or where the student failed to comprehend. They're indicating that their math and reading scores have risen considerably. It has a lot to do with—through a computer program—finding out where the youngsters were having problems and then tutoring to meet those problems. That's what testing should be all about.

That's my sermon for the day. You heard that many times in the National Council of Testing.

[The prepared statement of Hon. William F. Goodling follows:]

STATEMENT OF HON. WILLIAM F. GOODLING, A REPRESENTATIVE IN CONGRESS FROM THE STATE OF PENNSYLVANIA

Mr. Chairman. These hearings on educational standards and testing are critically important. If we are to develop a program and authorize these activities, we must do them with the best information available and with wide consultation.

I am pleased that Governor Romer is joining us this morning. The Governor is one of the greatest proponents of standards and testing and was the Chairman of the Council. It just wouldn't have been complete if we did not hear from him on this subject.

Once in a while, a report or study comes along that is both well timed and well done. The OTA report on testing is welcome on both these counts. The information on costs and current practices should give us the background we need, and the questions the report poses must be responded to before we move forward to authorize such a system.

I welcome our other witnesses, all of whom I know, and anticipate that they will provide for a very candid and lively discussion of these issues.

Chairman KILDEE. Thank you, Bill. Our first witness this morning is Dr. Michael Feuer, who is accompanied by Ms. Nancy Carson. Both are from the Office of Technology Assessment. If they would come forward. You may proceed.

STATEMENT OF MICHAEL J. FEUER, PH.D., SENIOR ANALYST AND PROJECT DIRECTOR, OFFICE OF TECHNOLOGY ASSESSMENT; ACCOMPANIED BY NANCY CARSON, OFFICE OF TECHNOLOGY ASSESSMENT

Mr. FEUER. Good morning, Mr. Chairman, Mr. Goodling. My name is Michael Feuer. I am the project director for OTA's study of educational testing. We're pleased that the release of our report is so timely, and we're glad to be here this morning to discuss the report with you.

Mr. Chairman, the American Educational System is unique in many ways. We were among the first in the world to experiment with the idea of public schooling, and we established the commitment to education for the masses well before any other country. Pluralism, diversity, local control—these are hallmarks of American democracy—have always distinguished the American educational experiment from those in other nations.

Standardized achievement tests, so-called because they are administered and scored under uniform conditions, have always played a central role in this educational experiment. Parents, legislators, and the general public have always sought assurances that

their children were getting a fair share of the advantages of education and that school resources were being managed wisely.

The idea that a relatively inexpensive tool could advance both the fairness and the efficiency of schooling has had obvious appeal to policymakers. With the invention of the multiple choice format, and later machine scoring, mass testing became increasingly economical.

Not surprisingly, testing has always been controversial. Precisely because tests have been promoted on both fairness and efficiency grounds, they have always been scrutinized on the questions of whether they do more to alleviate or aggravate social, economic, and educational disparities.

There's been substantial growth in standardized testing over the past three decades, as shown in the chart which we have displayed over here on the easel. This chart gives the percentage changes in revenues in constant dollars from sales of commercially produced standardized tests, tests used in elementary and secondary schools, and also the changes in public school enrollments over the 1960 base year.

The chart speaks for itself in describing the growth in testing over the period 1960 to 1989. Now this increased reliance on standardized testing continues a long tradition. It reflects widening public concern with the quality of schooling, a faith in the usefulness of quantitative data to aid in policy decisions, and, as in the past, it has engendered controversy.

Pressures to be accountable for our educational investments have led policymakers at all levels of government to demand more and more data. But many are questioning whether the information gained from tests is really what they need to make better decisions.

In this context, Mr. Chairman, the Federal role has been significant, requiring nationally comparable data as a basis for evaluating key Federal education programs such as Chapter One, has intensified the reliance on standardized tests. In fact, with respect to the chart, it's worth noting that the large increases in test sales seem to coincide with the advent of increased Federal and State accountability regulations.

Because the Federal role in testing has been so important in the past, congressional action now could have far-reaching impacts on the future of student testing as well. Our report raises several possibilities for Congress should it choose to foster an environment for continued testing innovation and reform.

For the purposes of today's hearing, though, I'd like to focus on two options that bear most directly on the current debate over national standards and national examinations. Before I do, though, it is important to set the stage with some of the principal messages of OTA's report. These points are discussed in some detail in the summary which you are releasing today and in much greater detail in the full report which will be out shortly.

Mr. Chairman, educational tests have three main purposes: to provide feedback to teachers and students that can help classroom learning; to evaluate the effectiveness of schools and school systems; and to inform important selection, placement, and credentialing decisions about individual students.

Using tests for these purposes is a delicate matter. Tests, after all, are only tools of estimation. Test results can vary for reasons that may have little to do with the student's real level of knowledge or achievement. Tests should be applied only to functions for which they are designed and validated. Yet, almost daily, we see examples where this simple guideline is ignored. Sadly, the misuse of tests has undermined their credibility even in the applications for which they are appropriate.

Attaching high stakes to test results is a particularly risky business. Labeling and ranking children or schools, pinning student promotions and graduations to test scores, issuing certificates of mastery based on test performance, have all heightened the controversy over testing.

In addition to raising questions about fair treatment of individual students, attaching significant rewards or sanctions to test results can compromise the validity of the information contained in the test results and mislead policymakers. In a word, high stakes testing can undermine both the fairness and deficiency properties of tests.

I'd like to emphasize that many standardized tests in use today could provide useful information to students, teachers, parents, and the general public. But their value is compromised by policies that encourage excessive focus on improved scores rather than increased learning.

It would be unfortunate if many of the exciting test innovations that now seem possible are similarly compromised by attaching high stakes to their results prematurely. What about the idea of using tests as a vehicle of education reform? This, too, is an idea that has been around since the first days of public schooling in the U.S.

Some new testing methods are especially appealing for this because they reflect important learning goals such as thinking and problem solving. But tests alone cannot stimulate reform without substantial preparation and support for teachers and schools.

In many States where new methods are being tried out, teachers are actively involved in all phases; definition of curricular and instructional goals, design of appropriate assessment tools, and the interpretation of results. On the other hand, a top-down approach, mandating tests and then expecting the whole system to respond, is a recipe for disappointment.

Finally, a word about costs. It is difficult to predict with any accuracy the costs of a major new testing program. However, the OTA report does contain an analysis of costs for existing tests in a large urban school district and also reviews data on the costs of designing and validating items for multiple choice tests as well as the costs of a major testing development effort in the military. Much of this information is summarized in my written statement and especially in Attachment A.

The main points I wish to make this morning in this regard are as follows: first, that decisions about testing need to include decisions about costs; second, that costs can be significant; and third, that the question of who bears the costs of testing ought to be a key issue in testing policy.

I would like to underscore one more point. Tests that appear very inexpensive can in fact be quite costly. Conversely, tests that appear costly can in fact be less burdensome than they really are. Let me try that one more time. Tests that appear inexpensive can be quite costly, in fact, and tests that appear expensive can in fact be less burdensome than they appear.

This depends to a large extent on how well the particular tests are integrated with instruction. With these as basic sorts of background, let me now proceed to two policy issues regarding national testing and the debate over standards. I'd like to begin by summarizing OTA's findings regarding the national assessment of educational progress.

The U.S. has had, since 1969, an invaluable program to gauge the general condition of education in the country as a whole. NAEP is the only regularly conducted national survey of educational achievement at the elementary, middle and high school levels. Its report cards provide information to researchers and policymakers concerned with the progress of student achievement over time.

NAEP has been a pioneer in test innovations, things like matrix sampling, broad-based consensus building processes, and the use of various types of open-ended or performance tasks. NAEP's biggest problem today, ironically, is surviving its own success.

Where it is largely due to its high credibility among policymakers and testing experts, NAEP is now being considered for new uses, uses that could ultimately jeopardize its original and still quite valid purpose.

OTA concludes that Congress should weigh very carefully any proposals to change NAEP and should focus rather on ways to retain and strengthen NAEP as a national indicator of progress.

Requiring NAEP to include more innovative tasks, restoring funding for NAEP testing and subject areas that have been omitted over the years such as art and music, finding better ways to communicate NAEP results to the general public, and expanding NAEP to assess knowledge in the adult nonschool population are some particularly promising avenues.

If Congress wishes to develop or support a new national testing system in which tests are administered to each child and used as a basis for important decisions about children in schools, OTA concludes that NAEP is simply not appropriate.

Let me turn now to the question of national testing. As you know, Mr. Chairman, the past year has witnessed a flurry of proposals to establish a national system of tests for elementary, middle, and high school students.

Momentum for these efforts has been fueled by reports warning of crisis in the economy and in the schools, by references to testing practices in other countries, by the national goals process, and most recently by the administration's American 2000 initiative.

Congress needs to clarify exactly what objectives are attached to proposals for national testing and how test instruments will be designed to accomplish those objectives. OTA finds that at present there is insufficient evidence to answer some of the key questions regarding national testing.

For example, one objective of a new national testing system seems to be to induce students and teachers to reach higher standards of performance. If so, it will most likely be a high stakes testing system. But what will happen to students, teachers, or schools who perform poorly? Will resources be provided to students who do not test well? Will the tests focus on broad domains of knowledge or force a narrowing of instruction toward specific topics covered on the tests?

Another question concerns the relationship between national tests and reform efforts already underway in many States and districts. Will this new system of examinations reinforce or hamper those efforts?

Third, if tests are to be somehow associated with national standards of achievement, who will participate in setting the standards? What evidence is there that existing or new tests can estimate progress towards reaching those standards? A fourth question is who will pay for the new tests, and who will govern their selection and certification as appropriate instruments?

Until these and other questions are addressed, OTA finds that placing a new test or system of tests into service at the national level could easily create new barriers to many educational reforms already underway and become the next object of concern and frustration in the American school system.

In closing, Mr. Chairman, I'd like to emphasize the importance of this period in American educational policy. We stand at the crossroads of genuine test reform. Greater understanding of how children think and learn, coupled with innovative technologies to assess their progress, could have profound affects on the quality and accessibility of education for all children.

To build on this opportunity requires careful understanding of the forms, functions and consequences of tests. Existing tests serve some functions well but become compromised when used excessively or inappropriately.

Although much of the interest in test innovation is spurred by a sense of urgency regarding the schools, education leaders should resist the temptation to plunge into major new testing programs using either existing or innovative test designs until the purposes and probable effects of those programs are clarified. New tests need a chance to evolve in an atmosphere conducive to careful study of their meaning and consequences.

Thank you for this opportunity to testify, and I look forward to answering questions that you might have.

[The prepared statement of Michael J. Feuer, Ph.D. follows:]

TESTING IN AMERICAN SCHOOLS: ASKING THE RIGHT QUESTIONS

STATEMENT OF DR. MICHAEL J. FEUER
Senior Analyst and Project Director
Science, Education and Transportation Program
Office of Technology Assessment
Congress of the United States

Testimony Before the
Subcommittee on Elementary, Secondary and Vocational Education
Committee on Education and Labor
United States House of Representatives

February 19, 1992

INTRODUCTION

Good morning, Mr. Chairman, Mr. Goodling, and members of the Subcommittee. My name is Michael Feuer. I am the project director for OTA's study of educational testing. We are pleased that the release of the report is so timely, and are glad to be here this morning to discuss the report with you.

OTA's study began 2 years ago, when the Committee asked us to examine a wide range of issues regarding educational testing. The Committee was particularly interested in the advent of new testing technologies. "Performance assessment," for example -- a term you have all heard much about -- was a relatively new concept then. So was the idea of using existing or new computer and video technologies, which were already establishing their presence in many schools for various testing purposes.

In addition to requesting information on the technological aspects of tests, though, the Committee emphasized the policy issues surrounding tests generally: How could existing or new tests be better applied to improve student learning, evaluate school effectiveness, and allocate scarce educational resources?

We analyzed both existing and emerging technologies for achievement testing, and the historical and current uses of tests to meet various educational and social objectives. In addition, the report draws some cautionary lessons from testing practices in other industrialized countries.

TESTING AND AMERICAN SCHOOLING

Mr. Chairman, the American educational system is unique in many ways. We were among the first in the world to experiment with the idea of public schooling, and established a commitment to education for the masses well before any other country. Pluralism, diversity, and local control -- hallmarks of American democracy -- have always distinguished the American educational experiment from those in other nations.

Standardized achievement tests -- so-called because they are administered and scored under uniform conditions -- have played a central role in this experiment. Parents, legislators, and the general public have always sought assurances that their children were getting a fair share of the advantages of education and that school resources were being managed wisely. The idea that a relatively inexpensive tool could advance both the fairness and the efficiency of schooling has had obvious appeal to policymakers. With the invention of the multiple-choice format and, later, machine scoring, mass testing became increasingly economical.

Not surprisingly, though, testing has always been controversial. Tests have always been scrutinized on the questions of whether they do more to alleviate or aggravate social, economic, and educational disparities.

The substantial growth in standardized testing over the past three decades then, continues a long tradition. It reflects widening public concern with the quality of schooling and a faith in the usefulness of quantitative data to aid in policy decisions. And it has engendered renewed controversy. Pressures to be accountable for our educational investments have led policymakers at all levels of government to demand more and more data; many are questioning whether the information gained from tests is what they really need to make better decisions. In this context the Federal role has been significant: requiring nationally comparable data as a basis for evaluating key Federal education programs has intensified our reliance on standardized tests.

Because the Federal role in testing has been so important in the past, congressional action could have far-reaching impacts on the future of student testing as well. Our report raises several possibilities for Congress, should it choose to foster an environment for continued innovation and reform. Today I will focus on two options that bear most directly on the current debate over national standards and examinations. Before I do, though, it is important to set the stage with some of the principal messages of OTA's report.

FUNCTIONS AND USES OF TESTS

Educational tests have three main purposes: to provide feedback to teachers and students that can help classroom learning, to evaluate the effectiveness of schools and school systems, and to inform important selection, placement, and credentialing decisions about individual students. These functions touch on basic principles of schooling in the United States: teachers and students are involved in an ongoing journey of learning that requires constant assessment and adjustment, schools and school systems must be accountable to the public that fund them, and important decisions about individuals must be fair and based on valid information about their performance.

But using tests for these purposes is a delicate matter. Tests, after all, are only tools of estimation, and test results can vary for reasons that may have little to do with a student's real level of knowledge or achievement. It should be obvious that tests should be applied to functions for which they are designed and validated. Yet almost daily we see examples where this simple guideline is ignored. Sadly, the historical misuse of tests has severely undermined their credibility even in applications for which they are well designed.

RAISING THE STAKES

Attaching "high stakes" to test results is a risky business. Labeling and ranking children or schools, pinning student promotions and graduation to test scores, or issuing certificates of mastery based on test performance have all heightened the controversy over testing. In addition to raising questions about fair treatment of individual students, attaching significant rewards or sanctions to test results can compromise the validity of the information contained in the test results and mislead policymakers.

I would like to emphasize that many standardized tests in use today are designed to provide useful information to students, teachers, parents, and the general public; but their value has been compromised by policies that encourage excessive focus on improved scores rather than increased learning. It would be unfortunate if many of the exciting test innovations that now seem possible were suffocated by hasty application of the tests to high-stakes decisions.

TESTING AND SCHOOL REFORM

Some new testing methods are especially appealing because they reflect important learning goals such as thinking and problem solving. But tests alone cannot stimulate reform without substantial preparation and support for teachers. In many States where new methods are being tried out, teachers are actively involved in all phases: definition of curricular and instructional goals, design of appropriate assessment tools, and interpretation of results. A "top-down" approach -- mandating tests and expecting the whole system to respond -- is a recipe for disappointment.

TESTING COSTS

Designing, developing, validating, implementing, and scoring tests are expensive processes.

To gain greater insight into the costs of existing tests, OTA obtained data on testing expenditures in a large, urban school district. We found that total direct expenditures -- materials, scoring, other contracted services, as well as time spent by teachers in test administration -- were roughly $13 per pupil per testing cycle. In addition, we explored the costs of teacher preparation time, i.e., the portion of teacher salaries accounted for by the time they spend preparing students for tests. These costs ranged from zero -- for teachers who said they spend no additional time preparing for tests -- to $100 per pupil -- for teachers who said they spend 3 weeks preparing before each testing period.

Even a modest innovation in testing -- revising and updating a conventional multiple-choice test -- is costly. It can take from 6 to 8 years to write new items, pilot test, and validate a major revision to a standardized achievement test battery that has been in use for many years.

Our analysis illustrates the effect on costs of defining to what extent testing is integral to instruction. The more a test is part of regular instruction, the lower its overall costs will appear, all else equal. (This point is explained in greater detail in attachment A, an excerpt from our report Summary.)

It is very difficult to obtain reliable data on the costs of new testing methods. Most performance assessment methods are too new to allow for accurate cost estimation. However, OTA did review some cost data from testing programs that have longer track records. (More data may soon become available from State efforts. In addition, the General Accounting Office has collected survey data on costs of a variety of testing methods.) First, we were able to locate cost data for essay examinations, which are the most well-understood form of performance assessment. In Ireland, for example, it costs about $135 to score a student's answers to five essay questions in each of four subjects. Similar data are now being reported for France, Switzerland, Italy, and England. In many countries where this type of examination is used, the costs are absorbed in the regular salaries of teachers, who are expected to do the scoring as part of their job descriptions.

Another insight to costs of test innovation comes from a recent effort in the armed services. It cost some $30 million, over 10 years, to develop a set of new performance measures assessing specific job-related skills in 30 well-defined military jobs. Although the results of this effort are informative, defining the content and developing measures for elementary and secondary school testing is much more complex than defining specific job performance outcomes for a number of jobs.

These findings make clear that the costs of testing -- the observed as well as the indirect costs -- will play a significant role in decisions about implementing new or existing tests.

With these points as background, let me now proceed to two policy issues concerning national testing.

POLICY ISSUES REGARDING NATIONAL TESTING

National Assessment of Educational Progress

The United States has had, since 1969, an invaluable program to gauge the general condition of education in the country as a whole. The National Assessment of Educational Progress, or NAEP, is the only regularly conducted national survey of educational achievement at the elementary, middle,

BEST COPY AVAILABLE

and high school levels. Its "Report Cards" provide information to researchers and policymakers concerned with the progress of student achievement over time. In addition, NAEP has been a pioneer in test innovations such as matrix-sampling, broad-based consensus building, and the use of various types of open-ended or performance tasks.

NAEP's biggest problem today is, ironically, surviving its own success. For it is largely due to its high credibility among policymakers and testing experts that NAEP is now being considered for new uses -- uses that could ultimately jeopardize its original and still valid purpose.

OTA concluded that Congress should weigh carefully any proposals to change NAEP, and should focus rather on ways to retain and strengthen it as a national indicator of educational progress. Requiring NAEP to include more innovative tasks, restoring funding for NAEP testing in subject areas such as art and music, finding better ways to communicate NAEP results to the general public, and expanding NAEP to assess knowledge in the adult nonschool population are particularly promising avenues.

However, if Congress wishes to develop or support a new national test -- to be administered to each child and used as a basis for important decisions about children and schools -- OTA concludes that NAEP is not appropriate.

Proposals for a System of National Examinations

The past year has witnessed a flurry of proposals to establish a national system of tests for elementary, middle, and high schools. Momentum for these efforts has been fueled by reports warning of crisis in the economy and the schools, by references to testing practices in other countries (see attachment B), by the National Goals process, and most recently by the Administration's America 2000 education initiative.

Congress needs to clarify exactly what objectives are attached to proposals for national testing and how test instruments will be designed to accomplish those objectives. OTA finds that at present there is insufficient evidence to answer key questions regarding national testing.

For example, one objective of a new national testing system is to induce students and teachers to reach higher standards of performance. But what will happen to students, teachers, or schools who perform poorly? Will resources be provided to students who do not test well? Will the tests focus on broad domains of knowledge or force a narrowing of instruction toward specific topics covered on the tests?

Another very important question concerns the relationship between national tests and reform efforts already under way in many States and school districts. Will the new system of examinations reinforce or hamper those efforts?

Third, if tests are to somehow be associated with national standards of achievement, who will participate in setting the standards? What evidence is there that existing or new tests can estimate progress toward reaching those standards?

Fourth, who will pay for the new tests? Who will govern the selection of tests and their certification as appropriate instruments?

Until these and other questions are addressed, OTA finds that placing a new test or system of tests into service at the national level could easily create new barriers to many educational reforms already under way and become the next object of concern and frustration within the American school system.

CONCLUDING REMARKS

In closing, Mr. Chairman, I would like to emphasize the importance of this period in American education policy. We stand at the crossroads of genuine test reform. Greater understanding of how children think and learn, coupled with innovative technologies to assess their progress, could have profound effects on the quality and accessibility of education for all children. To build on this opportunity requires careful understanding of the forms, functions, and consequences of tests. Existing tests serve some functions well, but become compromised when used inappropriately. Although much of the interest in test innovation is spurred by a sense of urgency regarding the status of the schools, education leaders should resist the temptation to plunge into major new testing programs -- using existing or innovative test designs -- until the purposes and probable effects of those programs are better clarified. New tests need a chance to evolve in an atmosphere conducive to careful study of their meaning and consequences.

Thank you, and I look forward to any questions you might have.

Attachment A

Box F—Costs of Standardized Testing in a Large Urban School District

Because testing policy decisions are still primarily made at the local and State levels, OTA has analyzed the kind of data on standardized testing costs that school authorities would likely include in their deliberations over testing reform. Data for this illustrative example were provided by the director of Testing and Evaluation in a large urban school district with 191,000 enrolled students, among whom 32 percent are in Chapter 1 programs. The district employs 12,000 teachers, including regular classroom and special teachers. The total 1990-91 district budget was $1.2 billion.

Approximately 140,000 students in grades kindergarten through 12 take tests, once a year in kindergarten and twice a year (fall and spring) in all other grades (absenteeism and student mobility account for the large number of untested students). During each test administration, students take separate tests in English, mathematics, social studies, and science. The tests typically consist of norm-referenced questions supplemented with locally developed criterion-referenced items. (In kindergarten, first, second, and third grades, criterion-referenced checklists filled out by teachers supplement the paper-and-pencil tests.) The tests are machine scored by the test publisher, who provides computer-generated score reports to district personnel. Tests are administered by 4,500 regular classroom teachers; there are no other special personnel involved, except for a small group of district staff who design the criterion-referenced items, manage the overall testing program, and conduct research based on test results.

Although the district purchases tests from a large commercial publishing company that has many school districts as customers, the cost figures discussed below are not necessarily representative of other school districts in the United States.

Materials and Services

In most years, the district purchases only a limited supply of test booklets, replacing the complete set only once every few years when they become damaged or when test items are revised. OTA computed average annual expenditures on test booklets based on test publishers' estimates that booklets are recycled typically once every 7 years. As shown in table F1, total annual outlays for the standardized testing program in 1990-91—including materials, contracted scoring and reporting services, and nonteaching personnel—were approximately $1.6 million, or $5.70 per student per test administration.[1]

Teacher Time

Based on the specified time allotments for the various tests in the various grades, and on conversations with district staff, OTA found that full-time teachers in the district spend roughly 2 percent of their annual work time in the *administration* of tests to students. The total salary cost to the district for teacher time spent administering tests was roughly $3.6 million for two testing administrations ($1.8 million per testing cycle).

Table F1—Outlays on Materials, Services, and Personnel

	Cost
Materials	
Contracted:	
Test booklets: new purchases plus annualized costs based on assumed 7-year cycle	$369,000
Practice books	49,400
Examiner manuals	26,200
Checklists and worksheets	100,600
Kindergarten program	33,300
Other:	
Kindergarten Chapter 1 tests	$3,000
Labels	1,200
Pencils	17,900
Answer sheets	23,000
Headers	2,700
Language battery	1,300
Special tests	14,100
Materials subtotal	$641,700
Services	
Contracted:	
Scoring	$175,600
Report generation	141,800
Collection	14,800
Scanning	146,500
Distribution	9,000
Services subtotal	$487,700
Nonteaching personnel:	
Assistant director	$58,200
Research manager	56,500
Research associates (2)	108,700
Research assistants (3)	127,800
Secretaries	56,500
Clerks	45,600
Nonteaching personnel subtotal	$453,300
Total	**$1,582,700**

SOURCE: Office of Technology Assessment, based on data supplied by a large urban school district. 1990-91 academic year.

[1]To understand how this district's cost of standardized testing compares with others, OTA looked at cost data from the November 1988, "Survey of Testing Practices and Issues," conducted by the National Association of Test Directors (NATD). The survey was sent to testing directors in approximately 125 school districts. For 38 districts providing their cost information, the average direct cost per student was $4.80 per year, slightly lower than the $5.70 per student in this example. Most of the districts responding to the NATD survey administer achievement tests only once a year, compared to OTA's example district, which tests twice a year in grades 1 to 12.

BEST COPY AVAILABLE

52-840 O - 92 - 4

Attachment A--Continued

Box F--Continued

In conversations with district teachers, OTA found that the time they spend in classroom *preparation* of students for the standardized tests varies from 0 to 3 weeks per testing administration. Some teachers claim they spend no time doing test preparation that is distinguishable from their regular classroom instruction; others use the standardized test as a final examination and offer students the benefit of lengthy in-class review time. OTA therefore estimated the salary costs for preparation time under three scenarios: 0, 1.5, and 3 weeks (per test). These estimates are summarized in table F2.

Table F2—Salary Costs of Teacher Time Spent on Testing, per Test Administration[a]

Test administration[b]	Test preparation	Total[c]
$1.8 million	0 weeks: 0	$1.8 million
	1.5 weeks:[d] $7.2 million	9.0 million
	3 weeks: $13.5 million	15.3 million

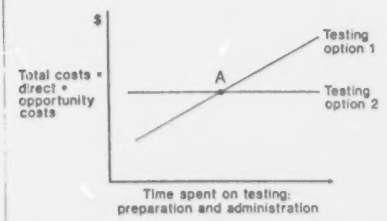
[a]Based on average salary of $40,500 per year.
[b]Based on an estimated 2 percent of total time spent on test administration for two testing periods.
[c]Based on 4,500 teachers.
[d]8 days.

SOURCE: Office of Technology Assessment, based on data supplied by a large urban school district, 1990-91 academic year.

Total Direct Costs

The total direct costs of testing can be computed by adding the expenditures on materials and services to the costs of teacher time for test preparation and administration. It is important to note, however, that this analysis does not account for the degree to which teacher time spent on testing is considered to be a necessary and well-integrated part of regular instruction. The importance of indirect or opportunity costs as it pertains to the analysis of testing costs is illustrated in box G.

Box G—Direct and Opportunity Costs of Testing



This figure illustrates the relationship between time spent on testing activity and the total costs of testing. Hypothetical test 1 is assumed to contribute little to classroom learning. It costs little in direct dollar outlays, but is dear in opportunity costs. Total costs begin relatively low but rise rapidly with time devoted by teachers and students to activities that take them away from instruction.

Hypothetical test 2, which is a useful instruction and learning tool, requires relatively high direct expenditures. But the opportunity costs of time devoted to testing are relatively low.

At point A, a school district would be indifferent between the two testing programs, if cost was the main consideration.

SOURCE: Office of Technology Assessment, 1992.

Attachment B

Box H—National Testing: Lessons From Overseas[1]

The American educational system has a traditional commitment to pluralism in the definition and control of curricula as well as the fair provision of educational opportunities to all children. Lessons from European and Asian examination systems, which have historically been geared principally toward selection, placement, and credentialing, need to be considered judiciously. OTA finds that the following factors should be considered when comparing examination systems overseas with those in the United States:

- Examination systems in almost every industrialized country are in flux. Changes over the past three decades have been quite radical in several countries. Nevertheless, there is still a relatively greater emphasis on tests used for selection, placement, and certification than in the United States.
- None of the countries studied by OTA has a single, centrally prescribed examination that is used for all purposes—classroom diagnosis, selection, and school accountability. Most examinations overseas are used today for certifying and sorting individual students, not for school or system accountability. Accountability in European countries is typically handled by a system of inspectors charged with overseeing school and examination quality. Some countries occasionally test samples of students to gauge nationwide achievement.
- External examinations before age 16 have all but disappeared from the countries in the European community. Primary certificates used to select students for secondary schools have been dropped as comprehensive education past the primary level has become available to all students.
- The United States is unique in the extensive use of standardized tests for young children. Current proposals for testing all American elementary school children with a commonly administered and graded examination would make the United States the only industrialized country to adopt this practice.
- There is great variation in the degree of central control over curriculum and testing in foreign countries. In some countries centrally prescribed curricula are used as a basis for required examinations (e.g., France, Italy, the Netherlands, Portugal, Sweden, Israel, Japan, China and, most recently, the United Kingdom). Other countries are more like the United States in the autonomy of States, provinces, or districts in setting curriculum and testing requirements (Australia, Canada, Germany, India, and Switzerland).
- Whether centrally developed or not, the examinations taken during and at the end of secondary school in other countries are not the same for all students. Syllabi in European countries determine subject-matter

[1]This draws on information from George Madaus, Boston College, and Thomas Kellaghan, St Patricks College, Dublin, "Student Examination Systems in the European Community: Lessons for the United States," OTA contractor report, June 1991.

Continued on next page

Attachment B–Continued

Box H—National Testing: Lessons From Overseas—Continued

content and examinations are based on them, organized in terms of traditional subject areas (language, mathematics, sciences, history, and geography) and, in some cases, levels at which the subject is studied (general or specialized). Even in European Community (EC) countries with a national system, the examinations are differentiated: all students do not take the same examination at the same time. The examinations may also be differentiated by locale (depending on the part of the country) or by track (there are high-level, low-level, and various curricular options).

- With differentiated examinations, multiple options give students on lower tracks the chance to choose lower level examinations. It appears, though, that these school-leaving examinations can discourage students who do not expect to do well from staying in school.
- In no other system do commercial test publishers play as central a role as they do in the United States. In EC and other industrialized nations, tests are typically established, tested, and scored by ministries of education, with some local delegation of authority. In Europe, Japan, and the U.S.S.R. the examinations have traditionally been dominated by and oriented toward the universities. In Europe, most examination systems are organized around a system of school inspectors, with quasi-governmental control through the establishment of local boards, or multiple boards in larger countries.
- Psychometrics does not play a significant role in the design or validation of tests in most European and Asian countries. Although issues of fairness and comparability are important, they are treated differently than in the United States.
- Teachers in other countries have considerable responsibility for administering and scoring examinations. In some countries (Germany, the U.S.S.R., and Sweden) they even grade their own students. Teacher contracts often include the expectation that they will develop or score examinations; they are sometimes offered extra summer pay to read examinations.
- Syllabi, topics, and even sample questions are widely publicized in advance of examinations, and it is not considered wrong to prepare explicitly for examinations. Annual publication of past examinations strongly influences instruction and learning.
- In European countries, the dominant form of examination is "essay on demand." These examinations require students to write essays of varying lengths in responses to short-answer or open-ended questions. Use of multiple-choice examinations is limited, except in Japan, where they are as prevalent as in the United States. Oral examinations are still common in some of the German *lander* and in foreign language testing in many countries. Performance assessments of other kinds (demonstrations and portfolios) are used for internal classroom assessment.

Chairman KILDEE. Thank you very much for your testimony. Let me start out with a question. Have you read or are you familiar with the National Council on Education Standards and Testing, their report that they issued in January?

Mr. FEUER. Yes. I'm familiar with it.

Chairman KILDEE. I thought you were. I just wanted to get that for the record. Let me ask you this question, then. Tests really indicate how one individual or how a school measures up to certain standards. You have a standard and you test to see how one measures up to that.

The national council on which Mr. Goodling and I served asserts the critical need for the highest possible national student performance standards. They call for national standards in which they measure outputs. But the Council puts down on the lower level the school delivery standards saying that they need not be national, but may be developed by a group of States.

If we have standards for students on one level, a higher level, but don't have standards for schools on that same level, does that not put a greater burden on the students rather than on the schools?

Just let me expatiate a bit on that. If you expect a student to score at a certain level in mathematics, at a certain grade level, the Council says okay, we're going to have national standards for that level but no national standards for the delivery system. Now, suppose we have in one school, school A, the Latin teacher, which I was, teaching calculus. I was great at Latin but not very good at calculus. But that happens.

Now, can't you almost predict to a great degree how the students are going to score in math if they have that commonly happening, where they're not having the teacher in that area of competence teaching? Can't you predict that those students will do probably less well on the test than those students who have had the teachers who are really trained to teach in that area? I'd like you to respond to that if you can gather the gist of what I am saying.

I know schools and I can tell where the teachers are being forced to teach outside their area of competence. That's just one area of standards. I know what the results might be. Could you comment on that?

Mr. FEUER. In an ideal world, we would be able to know exactly what the tests are intended to measure on the part of students, that is what outcomes we all want to be measuring. We'd also be able to explain differences in test scores by reference to the kinds of input measures that you're describing. So students in one school seem to do better than students in another school, and if that can be explained by the fact that Latin teachers are teaching calculus or calculus teachers are teaching Latin, that would explain, to some extent, the differences in the student's scores.

It seems to me, though, requiring standards for student performance is putting some of the onus on the students, and we are being more casual with the measures that could explain why students are performing at those levels.

I want to just add one further point and that is that I say that would be in an ideal world where we know exactly what we want

to be measuring in the way of student outcomes and where we have the tests that can actually do that measurement accurately.

Chairman KILDEE. Well, it seems like the word "casual" that you use is maybe an appropriate word. The National Council on Education Standards and Testing really emphasized the need for national standards for students and then became quite casual when it came to national standards for the schools.

Do you think they should have had a certain symmetry in establishing those standards?

Mr. FEUER. There has been a shift in attention and an interest toward outcome measures and away from input measures as a way to kind of hold schools and school systems accountable. So the emphasis on student performance is not surprising in that regard.

As far as the symmetry of the Council's work, I don't really think I'm in a good position to comment one way or the other.

Chairman KILDEE. You stated in your report that tests alone are usually poor vehicles to lead educational reform. If we did establish some standards for input, might that lessen or might that complement tests if we coupled that with some standards for input?

Mr. FEUER. I think I'd answer that by suggesting that rather than focus on trying to define standards for input that would be applicable everywhere, that we ought to look more carefully at the different approaches that are being tried in States and school districts which seem to have already a common objective, and that's to do a better job of educating all our children.

The fact that some school districts are experimenting with new approaches to curriculum definition and to instruction ought to be given a chance to flourish as a way of then explaining perhaps differences in the outcome measures. Setting standards prematurely could sort of force a homogenization of efforts that are still taking root in many places.

Chairman KILDEE. I think the Council is very skiddish about establishing a national curriculum. Aren't there certain basic things that should exist in educational institutions that would more likely lead to a better output? The one I used, of course, is very obvious, using a teacher that is not trained in that subject area. But that does happen. It happens a lot.

It happens very often in some of the poorer school districts also where, in competing with teachers, maybe they'll get a teacher and have them teach several courses as a means of saving some dollars in that school system. It seems to me that without going into a national curriculum, that there's a certain core of standards that a school could have without going into curriculum itself.

Could we separate standards from curriculum?

Mr. FEUER. Sure. I think we could, and I think we actually do. Most States have a pretty clearly defined set of curriculum objectives. People are working real hard to achieve those objectives. You know, part of the rhetoric that favors national——

Chairman KILDEE. It's just like producing a car. Everyone wants to have that perfect car. But if you have old equipment and lesser trained workers, you're going to have maybe a lesser quality car. Now, the curriculum at the auto company might be we're going to have the world standard car, but unless there's some standards placed into that, you're not going to achieve the results.

You wanted to comment on that?

Ms. CARSON. If I might. It just comes to mind that when we begin to talk about delivery standards and input standards, we're beginning to talk about resource allocation, and how to make standards real, and how to provide those things that we know schools and children need. So, if input standards become linked to resources and help to materialize the objectives, then I'm for input standards.

Chairman KILDEE. That's a very good summary. I think when you tie the two together, input standards to resources, what resources are used in that school system would make a difference, and that really would not be so much curriculum as to the resources you apply in that school. Very good. Thank you very much.

Mr. Goodling?

Mr. GOODLING. As many know, and I've said before, and probably will say at each one of these hearings, one of my chosen jobs serving on the Council was to slow down the train so the caboose didn't get ahead of the engine. If I'd had your study beforehand, I could have said it more eloquently than I did, I suppose.

I agree, and I've said many times, that NAEP should continue doing what they do well and should not get involved in whatever this new endeavor is. I also was concerned about attaching high stakes to national tests. I was also concerned for a while at the beginning because the Council seemed to be focusing on 3 years to come up with some test. I guess the one question I would ask is how many years do you think it would take to develop something that would be worth developing?

Mr. FEUER. Well, it takes from 6 to 8 years just to do a revision of a standardized achievement test battery that's already been in place for a long time. It took the National Council on Teachers of Mathematics about 6 years to define the standards. Now they are starting to work on the appropriate assessment tools to go with it.

In every place or every State or district, that has been experimenting with new approaches to testing, most places where they've been doing this, it begins with a very prolonged grassroots effort involving teachers, students, parents, community leaders, defining what the problems are and what the objectives ought to be. Once those curricular and instructional objectives seem to be clarified, they can then turn to the design of appropriate assessments.

So I couldn't go without saying it's a lengthy process. I couldn't pin down a number of years, but I know that in our work we spoke to a lot of teachers and a lot of parents and other people involved in the system.

There is a good bit of concern that rushing into something will give us either more of the kinds of tests that we have plenty of today or we'll kind of stymie these new and exciting efforts before they have a chance to be researched more fully.

Mr. GOODLING. I found your chart interesting. I'm only sorry that I didn't have money to buy stock in all of those companies. I have no other questions.

Chairman KILDEE. Thank you. I think we'll call upon people in the order of their appearance. Mr. Pastor?

Mr. PASTOR. Thank you, Mr. Chairman. I guess there's a sentence in your summary that I think is the key to this whole idea of

whether or not national tests increases level of students. But the sense is this, without adequate preparation and support, teachers cannot be expected to make effective use of any new instructional technology including new tests. I think that sentence is probably the best conclusion to report, "without adequate resources."

The example used by the Chairman of a teacher that has been prepared in Latin to teach calculus is a basic problem in our classroom. It doesn't matter what test you have and how well prepared the test is and what the standards are, if the teacher cannot provide that student just the substance of the curriculum, then we would have national tests in any number. But I agree with this sentence that that is the key conclusion to your report.

If the teacher is not prepared, if the teacher does not have the resources, then it doesn't matter what national test we have and what standards we have. It's not going to work. You may agree or disagree with me.

Mr. FEUER. I certainly agree with that.

Chairman KILDEE. Mr. Klug?

Mr. KLUG. No questions.

Chairman KILDEE. Mr. Henry?

Mr. HENRY. Thank you, Mr. Chairman. I am a bit concerned that we not get waylaid by the question you posed. I think your question about how you have a test for students, if you don't have tests for the criteria for the schools and how you hold a system accountable at the student level if you don't have criteria for the schools, misses the point.

Obviously it's related. You can't have students coming out without the supports for the schools. What is driving this issue of testing is the matter of simple public accountability. The public is not concerned as a consumer with an abstract assessment of some kind of test that assures them that their school has A, B, C, D, E, F programs and criteria. They are concerned about their kids not learning to read and write and compute.

It's as if we said we were going to solve the automobile problem in this country by ensuring we did everything possible we could by way of tax credits, by way of long-term R&D, by way of accelerated depreciation, by way of maybe a national health care system that relieves the manufacturers of the direct burden of health care costs, all these things and yet somehow the car coming out wasn't any better.

I just don't want us to try to escape from the fact that by virtually all of our measurements, our output, if you look at it in terms of a product level, has not been what we want it to be. I would caution all of us not to be beguiled into shifting the focus on process rather than product.

If you have a measure for the product, then you let the schools do different things that they may find appropriate in their communities, which addresses the issue of product output rather than dealing incessantly, as we have in education, with more and more process inputs.

Theoretically, that's what accreditation was all supposed to be about. Obviously, our accreditation agencies have failed us when they accredit schools that aren't doing the job. I'm talking about institutional accreditations. Obviously, our teacher accreditation

programs have failed us when in fact we readily admit there are many teachers out there who aren't, as was used in the example of the chairman, trained to be doing these things.

That's why the focus on the end is going to be the student. When I make, as a consumer, the choice as to whether my kid stays in a school system or not, it's not some abstract article in the Washington Post or my local newspaper that assures me that the school is making progress if I don't see my kid achieving what I want from my kid.

That's where the measurement has got to take place, it seems to me—I'm not taking out the process variable, but the whole point of national assessment and a national measure is looking at the end product. We're going to lose our discussion if we focus too much on the process, which is what I saw the chairman's comments moving us to, if that's appropriate.

Chairman KILDEE. I was moving toward that. You detected very well, except I do worry. I know you know national education because you've had a reputation both at the State level and here. In western Michigan, I would grant that most of the schools in your area would have good input standards. You'd have the right teachers teaching the right courses. You'd have good equipment, updated textbooks, not books that are 15 years old.

But there are certain areas in our country where you can find a concentration of some of these standards, some of the input standards that aren't really good, where they have, perhaps, teachers who are mismatched as to what their area of teaching should be. I've been in some schools where there were 20 year old textbooks, and the only computer may be the one the principal has in his office.

Now, there's going to be a difference, I think. This is, as was pointed out, the resources. I think if you emphasize maybe resources, some standards for resources of schools, that might be an appropriate——

Mr. HENRY. Well, to my colleague from Michigan, the new buzz word out there in education circles is this term "output-based education," trying to establish in your system your output goals and measuring them to see whether you got there. That's what national testing, or whatever we're going to call it, is about.

The fact of the matter is most our studies of outputs, to the extent that we have them, show strikingly causative relationships between inputs and outputs. There's such a wide variance. In fact, it's hard to tell what works and what doesn't when you don't have a uniform measurement scale.

So you've got to have the output-based measurement scale in the first place before you can even go back and do those things and say well, maybe the problem is they don't have enough computers, or they don't have enough continuing education for the teachers, or they don't have the support staff and psychological social services or whatever it may be.

What we've come to is the point saying you can't deal with inputs in a meaningful way until you have an accepted measure for the output. Accepted is the big issue. What's the accepted measure? Some kind of an objective. We've not had it. That's what you

need to do with these other things. You work backward right from that.

I just hope we don't lose sight of that. I share the chairman's concerns about those issues. He knows that. I've been involved in education a long time. But the push here is then to make it accountable to something. That's what we've got to have. Otherwise, the rest is just the same kind of education bologna we've had for 20 years.

Chairman KILDEE. Mrs. Unsoeld?

Mrs. UNSOELD. Nothing.

Chairman KILDEE. Mrs. Mink?

Mrs. MINK. Thank you very much, Mr. Chairman. This subject area of testing is very baffling. I'm very much interested in the national effort to try to find ways in which to evaluate our teaching, our schools systems, and the way it teaches in order to determine whether, in the process of teaching, we have failed to allow students to achieve at the level that we think they ought to be achieving, particularly in the language, the reading, science, math, those basic areas.

Yet, the dilemma is, as you've stated, how do you develop tests that can make accurate measurements, taking into consideration the wide diversity of our population, ethnic, economic, and so forth. Is there a valid public policy issue here which is drum-beating us to study this question about instituting national tests? If so, what is that public policy that brings this consideration to a national focus?

Mr. FEUER. I think there are certainly many issues that have converged on the idea of national testing as a way to formulate education policy. Certainly, there's an impression that America's economic strength is diminishing, that our competitiveness is weaker than it was, and that European and Asian countries have both superior schools systems and, to some extent, superior examination systems.

Mrs. MINK. But don't the schools already know individually how many students finish eighth grade and can't read, or don't have the competence in math when they finish high school, or don't know where the Capitol of the United States is on a map? I mean, aren't those things already known to the schools so that the establishment of a national test is not really giving the schools any more information than they already have, except for the fact that a national test will give a ranking? Is that the purpose? Is that the public policy?

Mr. FEUER. Well, that's certainly one of the recommendations we're making, is that the purpose of the policy be clarified.

Mrs. MINK. Is there a purpose to ranking?

Mr. FEUER. To ranking?

Mrs. MINK. Yes.

Mr. FEUER. Ranking is a fairly popular tradition in the United States. We rank a lot of things. We have a certain fascination with scoreboards and who's on top and who's in the middle and who's underneath.

Mrs. MINK. Well, haven't we got a whole bunch of those lists already? What have we accomplished by knowing you're at the bottom?

Mr. FEUER. We have a lot of data that could be used to make the kinds of comparisons that you're referring to.

Mrs. MINK. Well, why one more?

Mr. FEUER. Most comparisons don't end up surprising anybody.

Mrs. MINK. So what's the purpose of one more, something called a national test? To give it the good housekeeping seal of approval?

Mr. FEUER. I simply don't have a direct answer to that because our report is calling for a cautious approach to this whole idea. I think there's a panel that follows on our testimony which would probably be in a better position to answer the question as to what the purpose of a national ranking would be.

Mrs. MINK. Can you just sort of succinctly, not necessarily from your report but from a standpoint of a scholar in this area, run down a litany of what the public policy achievements might be if we adopted a national test? What could we learn from it that would be helpful in the course of trying to improve public education in America?

Mr. FEUER. I think we could continue to learn a great deal from national assessment of education if that assessment is based on, first of all, a range of indicators and not just scores from any single test or set of tests; and number two, if the data that are used for that kind of an assessment are kept as clear as possible from the kinds of interferences that arise with the sort of high stakes consequences.

That's why we're talking about keeping NAEP a strong resource. In the world of testing, NAEP is kind of up there as a national treasure. It would be a pity to undermine it, not that NAEP has no problems. I mean, NAEP is no perfect instrument either, but it's pretty good.

It's pretty good at doing the kind of thing that I think you are referring to, which is giving us a sense of how we are doing as a Nation. We could do a lot more to find out even at a deeper level where our weaknesses are and where our strengths are.

But you see, my concern is that if we either use NAEP or some other thing in which everybody will focus their attention on getting those scores up so that they look better in the next round, or in the next newspaper ranking, or in the next set of real estate ads that come out, that we will end up, first of all, being very unsure as to just what kind of progress we're making.

We have learned something about the effects of these high stakes tests over the last 10 or 15 years. One of the things we learn is that policymakers are not getting particularly reliable information. NAEP, on the other hand, has been very strong because it's been there as a kind of bulwark against all of the other kinds of less reliable data that are out there. It wasn't succinct, but getting there.

Mrs. MINK. If the purpose of the tests are to try to determine what the deficits are in public education in any given area, isn't it also essential, having discovered what the deficits are, that somebody is prepared to fill it?

Mr. FEUER. Well, that's exactly one of the questions that we're suggesting that Congress ask and insist on getting an answer to before it embarks on a new system of tests.

One of the ironies of testing is that we seem to have a rather widespread sense that there are problems in the schools. Much of

that comes from the testing that we do. To do more testing is to suggest that the information that we have is not valid. It's now a question of using it and allocating resources appropriately.

Mrs. MINK. Thank you very much, Mr. Chairman.

Chairman KILDEE. You're welcome. Thank you.

Mr. Gunderson?

Mr. GUNDERSON. Thank you, Mr. Chairman. Probably I'm missing something from your report and from the questioning that has been going on. But isn't part of what we're trying to do here to recognize the mobility of the American family and therefore the American student?

What is producing the incentive for testing is not so much even educational reform as it is a desire by families, and parents in particular, to be able to get some kind of a comparative analysis of a school in Wisconsin with a school in California they just left or the school in Florida they're just moving to. Parents simply need some idea of their student's level compared to the new school they're moving into?

Am I all mixed up on this? When I talk to families, that's what I'm hearing.

Ms. CARSON. Do you want me to try that?

Mr. GUNDERSON. You can try it, sure.

Ms. CARSON. Let me try.

Mr. GUNDERSON. Try to educate me.

Ms. CARSON. Certainly, we are still a mobile society. When parents move, one of their main concerns is finding a good school environment for their child. It is unfortunate that we don't have better signals about what constitutes a good school environment, because often students are forced back or do look at test outcome, test scores.

These are not very good indicators of schools, of delivery systems, of what is going on in the school. Most of the scores that we have available to parents tend much more to reflect how the children come to school, their backgrounds, other things as well as education.

I think certainly part of the desire to set standards and part of the impetus between the discussion of national standards is a belief that after all, we are a mobile country. People do move, and you shouldn't have something very different for a child in this State than in that State. But most States, in fact, do have standards. School districts do have curriculum. A parent could discover by visiting schools what is going on in that school.

To try to answer more succinctly, it would be nice if there were a simple way to tell what the best schools are. The kind of score outcomes that we have are not very good indicators, and there's no real good reason to think that putting in place a broader kind of examination would lead to, really, better information for parents on schools.

Mr. GUNDERSON. I disagree with you totally. But let me use an example. We can have a standard that says we have to have four semesters of math in high school. That doesn't mean the kid is going to learn anything in math. So you need much more than a standard in order to allow——

I think the problem with this whole debate on testing is we assume that we're going to use testing solely for the purpose of making accusations about someone being in a bad school, or a bad teacher, or having a bad program. I look at it just the opposite. It's like going to the doctor to have a physical. The doctor doesn't tell you you're a bad person. He said look, you have too much cholesterol. You have too much fat. You're smoking too much. You need to change your diet in this direction or that direction. I mean, isn't that what we're talking about here? It's not a beat up on somebody; it is how do we use different mechanisms to determine how to improve what you have?

Mr. FEUER. I think the key there, and in your language, it's different mechanisms. I would agree with you that to some extent, test scores could be one of many indicators that parents could look at to get an idea of what's going on in a school district or in a place where they're planning to move to.

But I want to just add to that that one of the ironies is that we are at the same time trying to encourage a great deal of diversity in our schools. The president's initiative has this very interesting idea of initiating all kinds of experiments in delivery of curricula and instruction.

Nationalizing the testing and assessment function and pushing toward a common set of outcomes it's not entirely clear that those two reconcile with one another. So it's certainly laudable to have a system in place that would allow parents to have more accurate information about what's going on in schools. That really ought to be based on multiple indicators.

Some parents frankly want a school that has a very good athletic problem, and others want a school that has a very good chemistry program. I'm not sure how one combines athletics and chemistry in a test score.

Mr. GUNDERSON. I agree with you totally, but I think the misperception here is there is going to be one uniform national test. I don't know of anybody who has advocated that. Am I missing something?

Mr. FEUER. No. I think that's right. We're certainly not advocating a single national test either.

Mr. GUNDERSON. Thank you.

Chairman KILDEE. On that, I think that standards and testing somehow must go together. You test something. If you don't have standards—I can go into certain schools that I visited here in America and walk out of that school and predict what the test results are going to be just by looking, as you mentioned, to the resources that are there.

I can predict that. I don't really need a great computer, a written test. I can just predict what's going to happen when they give the test in that school by looking at resources. So standards and testing do go together. I guess how we put those together is one of the things that this committee has to deal with very carefully.

Dr. Feuer and Ms. Carson have been very helpful in this. I, myself, may want to submit some additional questions to you in writing because we do have another panel. I know two of those people in the panel have to leave shortly.

Mr. Martinez?

Mr. MARTINEZ. Thank you, Mr. Chairman. I'll try to be quick here. I can understand if Mr. Gunderson is confused. I agree in one sense with what he says, but the question is, do you test to compare or do you test to know what you need to do?

You gave the analogy of a doctor. Now, I'm in total agreement when you say you test to know what to do, because the tests I've seen in my life have not been that at all. They've been just to compare to see where the bad people are. We don't need more tests to know where the bad schools are and where the pupils are not learning.

Too many times tests are misused. When I went to school, you might say I had a language barrier problem because our family did not speak English at home. So I was not put in an "English proficiency" class; I was put in a "speech correction" class. All the other students in there were people who stuttered or had some other speech impediment.

My time in that program was wasted. Meanwhile, being placed in that program put me behind because I wasn't learning. So, as far as I'm concerned, if you're going to test, you ought to test to know what to do. Now, if you're talking about testing against other countries and what they're doing, that may be all well and good just to tell yourself what you're competing with, as in the analogy you made about a pole vaulter.

But the truth of the matter is that still doesn't get back to helping the individual kid. Mr. Henry was talking about input and output. I'm not sure which he considers input and which he considers output. But if it was the way I understood, I think he's got to understand that before you can have output, productive output, you better have darn good input, which means you better set standards for schools and teachers and give them something to measure up to.

Then you can test the child to find out where he is—if what they are doing—if they are meeting the standards because the result of their meeting the standards will be that that young child is learning.

You also say in your testimony that we're not teaching students the things they need to know to compete effectively. You're absolutely right. But one of the things we're not teaching, and you do talk about it on the second page of your testimony, is how to think.

You know, Terrel Bell sat in one of these rooms and testified before us. He said something that really rang home. He said there's nothing so rewarding to a good teacher as to see the light go on in that young person's mind when they learn that they can learn.

I remember a sixth grade teacher, Ms. Kasems, who knocked me around like you wouldn't believe until she finally got me to learn that I could learn. I thought she was the meanest thing in the world until when I started crowing about how I learned, how I could do it, now I understood it, I looked up at her and she was smiling from ear to ear.

I think I thought, "well, she's not mean at all; she's just concerned about me." Sometimes you don't find those teachers. But if you have standards for teachers, and you have instructions for teachers, and you have standards for schools, then I think you're

going to start moving toward being able to test the kid to see what the kid has attained and where that kid is lacking so that he can move up from there.

That's the only kind of testing that I'm really interested in. I don't believe that one single national test can do it. I do think that some standards, national standards, sure as heck could get us going in that direction.

You say also we talk about the U.S. as the only industrialized Nation that does not have national testing. Tell me something about the achievements of that national testing in those other industrialized countries?

Mr. FEUER. Well, we don't say that the U.S. is the only industrialized country that has no national test. I want to try to clarify this. We have an attachment to the testimony which is an excerpt from the summary of the report. It tries to lay out some of the comparisons between testing policies in other countries and in the U.S.

Just to give you a smattering of some of what we found in our analysis of how other countries test compared to the U.S., for one thing, most testing systems in Europe and in Asia are in transition. Just as in the U.S., there's a big debate about how and what and when to test, this is something that's going on in most countries of the world.

Second, perhaps more important, is that no country that we looked at has what would be called a national examination for children before the age of 16. In countries where they are doing some kind of national testing, it is still largely focused on the selection and placement and credentialing function of tests, that is the gatekeeping function.

They don't do much testing there for accountability. They don't do much testing on a national basis there for the kind of classroom diagnostics that I think you were referring to earlier. Within and among the European countries, there's a considerable amount of variation in the degree to which testing is actually centralized.

In Germany, for example, there's a lot of autonomy in what are equivalent to States and regions. In a country like France, it's much more centralized. With respect to test format, Japan uses more multiple choice testing than we do. Germany uses more oral and written testing.

So, to the extent that we want to sort of imitate our fiercest competitors, we should choose which competitor we want to imitate. We could end up doing more multiple choice testing if Japan were the model and less if Germany were the model.

A lot of these comparisons need to be viewed in the light of some of these subtleties and complexities of how testing is organized. I guess the overarching difference is that schooling is quite different in most countries from schooling in the U.S. There are broader differences in the objectives of schooling and the reach of schooling, the accessibility. This is undoubtedly reflected in testing policy.

Mr. MARTINEZ. I think what I hear you saying is that we can look at all these things and gain some knowledge from them, but we shouldn't be trying to mimic or duplicate what they're doing. We ought to figure out what's wrong with our system and address the needs of our system.

Mr. FEUER. I think we can learn from overseas. We have to learn selectively and apply that which fits the best to our situation.

Mr. MARTINEZ. Thank you. Thank you, Mr. Chairman.

Chairman KILDEE. Mrs. Lowey?

Mrs. LOWEY. I have no questions.

Chairman KILDEE. If there are no additional questions, I may submit some questions in writing.

Just before you leave, I read the New York Times this morning about Mr. Buchanan's victory in New Hampshire. After that, I read another story. Interesting, I have written down on my paper now standards/resources and then testing below that. It states here that just when politicians are calling for a revival in American education, the recession is leaving principals and teachers across the country with less money for more students.

It goes on to say in many cases, rising school enrollments, local reluctance to raise taxes and State fiscal problems meant that schools actually have less money than before. That's one thing that I—all through this discussion this past year have tried to do—run these two on parallel paths, standards in testing and make sure that we don't put one ahead of the other. Maybe one should go ahead of the other. Maybe we should set the standards first and then have the testing; right?

I really appreciate your testimony here this morning. It was very helpful. Your report is extremely helpful. Thank you very much.

Now our panel will consist of Honorable Roy Romer, a good friend of mine—we worked very closely in the vineyard together on this—the Governor of the State of Colorado and the co-chair and workaholic, I will say, on the National Council on Education, Standards and Testing, who kept us all moving and working very, very hard.

Also, a member of that council, is Keith Geiger, president of the National Education Association and a resident, when I was growing up, or teaching I should say, in Michigan; Bella Rosenberg, Assistant to the President Al Shanker, American Federation of Teachers, AFL–CIO; Daniel M. Koretz, Ph.D., Senior Social Scientist, the RAND Corporation; and Michael H. Kean, Dr. Kean, Ph.D., Chair, Test Committee, Association of American Publishers and Vice President, Public and Governmental Affairs, CTB MacMillan/McGraw-Hill.

We'll lead off with Governor Romer, followed by Mr. Geiger, because I do know both of those have other appointments, pressing appointments they must make.

Governor, it's always good to be with you.

STATEMENTS OF HON. ROY ROMER, GOVERNOR, STATE OF COLORADO, CO-CHAIR, NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING; KEITH GEIGER, PRESIDENT, NATIONAL EDUCATION ASSOCIATION AND MEMBER, NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING; BELLA ROSENBERG, ASSISTANT TO THE PRESIDENT, AMERICAN FEDERATION OF TEACHERS, AFL-CIO; DANIEL M. KORETZ, PH.D., SENIOR SOCIAL SCIENTIST, RAND CORPORATION; AND MICHAEL H. KEAN, PH.D., CHAIR, TEST COMMITTEE, ASSOCIATION OF AMERICAN PUBLISHERS AND VICE PRESIDENT, PUBLIC AND GOVERNMENTAL AFFAIRS, CTB MACMILLAN/McGRAW-HILL

Governor ROMER. Thank you very much, Congressman. It's good to be back with you.

I obviously take this issue very seriously. I became very uncomfortable about the last bit of testimony. We really need to clarify what's on the table. You know, Congresswoman Mink, none of us are advocating a national exam. I can't remember in the panel, maybe in one or two persons, it's not an issue. I really want us to focus on what the issue is on the table. I was disappointed, quite frankly, that the person who preceded me allowed that to be the issue. Maybe it was an issue in his report, but it was not an issue in the measure which I think is before us in the panel.

Let me put this in context. I want to hit, if I can, very quickly, all the issues that I see on the table. When I gave my State of the State message, the next morning the headline of the Rocky Mountain news in Denver was "Governor Threatens: Raise Taxes or Else." Now that's an interesting headline for a politician, but it was very accurate.

I said to my legislature, you either do it right or we're not going home. I'm not going to sign the budget bill. Now, I'm trying to raise funds for education. So I want you to know, all of you who are interested in equity, I'm here. I'm with you on equity. We've got to do this job better.

The next point, one of the most important arguments against more money and education in Colorado is people feel that we're not getting our money's worth. We're not getting what we're paying for. Now, that's the issue. If I could convince people in Colorado that we're really doing a good job, they'd raise the money for it. I know that. I know the people of Colorado. So, I've got a real problem of convincing people in Colorado that we are doing the right kind of jobs in schools.

Now, one of the real problems we have out there is that there is another point of view, and it's almost contradictory. That is, a whole lot of people feel like what we're doing is good enough in terms of the standards. This issue of standards and testing, let me just talk about them separately and then together.

Why do we need to be concerned about standards? First of all, in math, we don't know what real math is in Colorado. We don't know what math power should be. Now we think we do, but we don't. So the first thing that we've got to do is to communicate to the people of this country that we are reaching for a standard that is not going to be successful for us.

The best way I can illustrate it is if I were at a blackboard, I would draw a pole vault. We've been vaulting out there at 15 feet in Iowa, Colorado, for years. Somebody come along and said wait a minute, the rest of the world is doing 19 feet. If we found that out, we'd sure change our way of vaulting.

What we need to do in education is send a message clearly as to what is the level of achievement we need to reach. What is an educated person? If we don't do that one, then I don't think we can do the whole list of other educational reforms. So I would make a very strong plea to you and Congress to help us get this job done.

Now, let me take the consequence if you don't, if you don't pass this bill. Let me just lay out what history will do. We'll have to do it anyway. In Colorado, I have a committee in which the president of the Senate, who is a Republican, the Majority Leader of the House, the Speaker of the House, is on it. It's arriving at standards for Colorado.

What do we want? We want to look somewhere in the Nation and say give us the best that you can about what it is that we can model ours after.

So, the other thing I want to underscore, this is voluntary. This is not mandatory. You know, we're really just asking collectively can we arrive at something at a national level that could assist our individual States to arrive at the right standards.

We need, in Colorado, whether you help us or not, we're going to arrive at what we hope are world class standards in math and English and science. If we don't get that on the table first, we never can do the other reforms to teach to it.

Let me then turn to assessment. I think the issues here are one, is it right to assess before we do all the other kinds of reform? Namely, I think one of you has just recently spoke and said that it isn't really fair to put higher standards and assessment on kids unless we can do a better job of preparing teachers and funding schools.

We need to do them simultaneously because if we do not adequately and accurately access what we're doing, I don't believe we will really know what is the level that we're reaching for in education. High stakes are a problem under our existing methods or under new and better assessments.

So, in a very brief way, what I'm appealing for is your assistance in passing this bill to create a mechanism where we will not mandate anything, where we will simply assist States like Colorado to arrive at the best standards, and they will be voluntary.

Secondly, in assessment, we're asking absolutely not for the development of a national test. We're not asking for the development of additional tests, more testing. What we're asking for is let's go and develop tests that really do what it is we think ought to be done.

Let me give you a simple illustration out of my own life. Standardized testing, I think, has some use, but we have grossly exaggerated the use of standardized multiple choice tests. We need to develop enriched kinds of testing so that we can really know what it is that we're after.

I'm a pilot. When I am in a test flying over the City of Washington, DC, as a student, one of the tests that I would be given is my

instructor would cut the engine. He'd stop it. He's say, "Romer, land the plane." Now, I would remember all those isolated bits of information such as here's the fuel system.

This is the ignition system and how it works. Here's how the wind goes over the wing and gives you a lift or so much glide. This is the area in which you need to land. You need to land into the wind. This is your breaking power.

Now, you can give me a multiple choice standardized test on all those isolated bits of information, decontextualization. That's the way in which we do this. Decomposability, those are the standard words of psychometrics.

But I could regurgitate all those isolated bits of information and still kill myself trying to land that plane because life is not like that. Life is like this, when that engine stops and that prop stares me in the face, what do I do? I have to use my mind, Congressman. I have to think. I have to organize that information. I have to sort it, collect it. I have to reason from it. I have to act upon it and communicate from it. Now, that's what life is like.

Now, that is a performance test. Now, what does that test do? First of all, it's very educational. Secondly, it points me to truly what is an educated person. An educated person is not someone who takes isolated bits of information and regurgitates them in some standardized test. An educated person is one who has the ability to utilize information, to organize it, to reason from it, to act upon it and to communicate based upon it.

Now, what I'm asking, then, is your assistance so that we can help the States of this country begin in their own pluralistic way to develop better forms of assessment. Assessment is absolutely critically linked to standards. Assessment really should do many things for a student. First of all, it should aim the student as to what it is that an educated person is.

I, therefore, because we need to keep this short, would make a strong appeal to you, don't let this bill bog down. We need it.

I want to talk about time lines. I got extremely uncomfortable when the prior speaker—and I respect him—Dr. Feuer, reported, and he's accurate in his report, that history has required 8 years to come up with the national council of standards in math. It takes 5 or 6 years to do a new standardized test.

We do not have the luxury of taking that long. If we can go into the Persian Gulf in 6 weeks with Desert Storm, we sure as hell can develop tests quicker than 6 years. It can be done. We've got bad habits here. We ought not be fearful of rushing too fast.

I look at every youngster in the schools of Colorado. Every year that goes by that we put resources into educating those youngsters—and we're not telling them the truth, which is there is going to be a high stakes test for you someday, and that is you're going to try to get a job.

If you have to be able to use math power realistically and accurately to get a job, we need to tell them early on what that is so that they can prepare for it. So, in summary, we need this capacity to develop standards to be offered voluntarily to States. We need this capacity to encourage the development of pluralistic methods of assessment of a better quality not to be mandated, not to be a

single national test, and not to be used as high stakes in the wrong way.

Let me finally close on that. High stakes, who makes the decision of a high stake? It's local, absolutely local. It's going to be continued to be local. So, what we need to do is to help folks locally. Let's take a fourth grader in math.

If we gave them a really good authentic test in math, and the youngster is not up to where the youngster ought to be, what should be the stake? I think the stake ought to be that the system should respond and say Joe or Mary, we're going to give you more time to learn and better ability to learn.

There was one other issue I wanted to get to before I close, and that is system standards as compared to student, content, and performance standards. There is a natural division between these, and I'd like to draw that. I think it's very appropriate for us to arrive at standards consensually and collectively nationally as to what the content ought to be and how good is good enough performance ought to be.

But I think it's very logical to leave to local States and districts various ways of arriving at that result. It would be dangerous, in my judgment, if we tried to proscribe at a national level delivery standards for individual schools or States, because pluralism of approach is a healthy thing.

Now, I think it is correct that as we develop student's performance standards that we should also say it's not fair to put the standard out unless we held the institution accountable to give the student the opportunity to learn. That is correct.

But we ought to allow the criteria for delivery standards to be developed, I believe, State by State in a collective action because we avoid something that is real in American politics, and that is there is a very strong history of education being local and that we do not find a way to proscribe a standardized way of going at education.

So, I think there's a legitimate distinction between national standards of performance that we want to achieve, but a variety of ways to achieve them, which are a variety of delivery systems. But we ought to say there ought to be a standard locally so that you can get at that issue.

Mr. Chairman, those are my comments, and I would be happy to participate in any questions.

[The prepared statement of Hon. Roy Romer follows:]

STATEMENT OF HON. ROY ROMER, GOVERNOR OF COLORADO

I am pleased to testify for the subcommittee about the report of the National Council on Education Standards and Testing (NCEST).

I am here today to impress upon you one simple message. This Nation is not teaching its students the things they need to know to compete effectively in the world economy. Our standards are too low, and the ways we go about measuring progress are inadequate.

It's useful to think of this issue in terms of a pole vault competition. If our goal is to win, then we ought to find out how high the competition is vaulting and that's where we should set our mark. If we set the bar at 16 feet, but the rest of the world sets it at 19 feet, we lose.

The analogy is appropriate. We have said we want world class education standards so that our kids are able to compete in the world economy. But we've never defined what those standards are or how we know when we get there.

In spite of all the time and resources devoted to testing in this country, we still are not able to tell exactly how far and in what ways students' performance falls short of world class standards. That's because we have traditionally measured success by our students' performance against one another or against multiple choice tests that focus on basic skills and small bits of information.

Neither measure is relevant to our competition—the world—and both can end up leading to a false sense of satisfaction.

Polls show that Americans want a world class education system. But all available evidence suggests that we are falling far short of that goal. And polls reflect a huge gap between belief and action, highlighting a dangerous sense of self-satisfaction.

The results of a recent Harris Poll should alarm the business and education communities alike. The study identified 15 key attributes that are basic objectives of secondary education and which are necessary for students to perform well on the job or in higher education. The attributes included:

—the basic skills of reading, writing, communication, and computation,
—personal work habits,
—the ability to work with others, and
—problem solving.

Representatives from higher education, employers, recent students, and their parents were all asked to evaluate the preparedness of students on these key attributes on a common scale.

The results were alarming in two respects. First, the very low ranks given by both employers and those in charge of higher education indicate that students are graduating without the skills and abilities they need to perform well in higher education or on the job. Significant remediation efforts are required.

Even more alarming is the dramatic gap between the very low rating employers and higher education officials give to student competency and skills and the favorable assessment of students and their parents. In effect, we have found out that while we have been teaching our students to aim for a 16-foot pole vault, the other countries of the world have been aiming for 19 feet. Parents and students have been satisfied with the 16-foot performance because the students have only been compared to each other, not to world class standards.

I served as the first chairman of the National Education Goals Panel. The purpose of the National Goals is to capture the attention and resolve of Americans to restructure our schools and radically increase our expectations for student performance. They cover the full range of learning from readiness for school to lifelong learning. They are consciously outcome-oriented. The report card is not designed to lay blame or indict. It is not another Wall Chart. It is intended to provide us with accurate feedback in a consistent way over time to improve our performance.

The Goals Panel members found that, particularly for goals three and four on academic performance, we don't have ways to authentically assess what children know and are able to do. Moreover, we have no national standards of performance against which to assess how well we are doing in achieving the goals as a Nation.

We don't know how good is good enough. We don't know how high our competition is vaulting. We do know from international assessments that our students do not compare well to other developed nations. The standards we use are far lower than other countries.

The Panel firmly believes that before we can seriously improve our educational achievement, we need to understand how short of the mark we are. I compare this to the use of a CAT scan in medicine. We first need to decide what we mean by "health." We need a standard. A CAT scan can tell us where we stand with regard to our standard of health. Then we can determine what needs to happen to get us to that goal.

The National Council on Standards and Testing was created to address this issue. The charge of the Council was to assess the desirability and feasibility of national standards and a system of national assessments tied to national standards.

As co-chair of the Council with Governor Campbell of South Carolina, we achieved an extraordinary, bipartisan and broad-based commitment to developing standards and new forms of assessments, first in the five subject areas in the National Goals and later in other areas such as the fine arts and citizenship. Important workplace skills and knowledge such as the SCANS commission report (the Secretary's Commission on Achieving Necessary Skills) describes should be included.

The Council just finished its work and issued our report, *Raising Standards for American Education*. In it, we provide the argument for developing national standards and a system of national assessments. Standards would let us know what we expect our students to know and be able to do at various points in their education.

They would enable students, parents, teachers, and the public to know where students are relative to a standard rather than relative to each other.

In the absence of well-defined and demanding standards, the Nation has gravitated toward de facto minimum expectations. High national standards tied to a system of assessments can create high expectations about what all students should know and be able to do. They can help to better target our educational resources.

National standards are not a national curriculum. We advocate standards with the following characteristics:

—reflect high expectations, not minimum competencies,
—provide focus and direction, not become a national curriculum,
—national, not Federal,
—voluntary, not mandated by the Federal Government,
—dynamic, not static.

Since tests demonstrate and determine what is important for students to learn, assessments should be developed that embody the new high standards. Much current testing reinforces the emphasis on low-level skills and bits of information taught in our schools. To make high standards meaningful, it is important that the Nation be able to measure progress toward them with new forms of assessments—tests that are worth teaching to and are designed to guide curriculum development.

The Council came out decidedly against a single, national achievement test in favor of a system of assessments to accomplish multiple purposes. In addition, the Council envisions new forms of assessments. Current tests are neither oriented toward the goals nor supportive of the curriculum and instructional changes needed to achieve them. New forms of assessment are needed that will drive changes in what happens in classrooms.

The considerable resources and effort the Nation currently expends on the patchwork of tests should be redirected toward the development of a new system of assessments. Assessments should be state-of-the-art, building on the best tests available and incorporating new methods such as performance assessment.

The national system of assessments should have two components:

—individual student assessments, and
—large-scale sample assessments such as the National Assessment of Educational Progress.

The purpose of having two components is that no single assessment can be used for multiple functions. We need to continue the matrix sample approach used by NAEP to assess the progress of the Nation in achieving the goals. NAEP should be tied to high national standards and should incorporate new methods of assessment.

New break-the-mold individual assessments must be developed for instructional use. These assessments must also be tied to national standards, to provide information to teachers, parents and students. These assessments should replace current tests rather than be in addition to them.

The key features of both components would be:

—alignment with high national standards,
—the capacity to produce useful, comparable results,
—multiple methods of measuring progress,
—voluntary, not mandated by the Federal Government,
—developmental, not static.

In order to effectively accomplish the Council's recommendations, we have reached agreement between the National Governors' Association and the administration that a new coordinating structure will be put into place. Of great importance, the National Goals Panel will be reconfigured to be politically balanced and to include Congress.

A permanent National Education Council on Standards and Testing would be established, whose members would be appointed by the Goals Panel. The Panel and the Council would jointly work to certify that content and performance standards are at world class levels and to set the criteria for a national system of assessments. The permanent Council will be the technical arm of the Panel to carry out the work of coordinating the development of new standards and certifying assessments.

High national standards and a system of assessments, while critically important, are not panaceas for the Nation's educational problems. Other elements of reform are necessary to enable us to reach high standards of achievement, including professional development, technology, incentives, early intervention strategies, and reducing health and social barriers to learning. High national standards and a system of assessments of the standards are the next step in revitalizing American education.

We know that local control and local decisions are critical for the accomplishment of fundamental changes. The establishment of national standards and assessments

will have to draw strength from a variety of approaches, including current work in the States. We need to capitalize on the knowledge and work being done around the country. We need to promote cooperation among States to develop appropriate means of assessment and establish procedures for comparing different kinds of information across States.

I want to specifically address some concerns and questions that I have heard about the Council's recommendations. The following four issues have come to my attention:

—The importance of the link between standards and assessments,
—Some misunderstandings about the NCEST report,
—My position on equity and delivery standards, and
—The resources needed to implement the Council's recommendations.

1. Why do we need both national standards and a national system of assessments?

As I described earlier, standards and assessments lie at the heart of significant educational reform. We must have a vision of what we expect students to accomplish and ways to assess their attainment of what is important to know and be able to do. Current forms of testing constrain our ability to change what we expect from our students. We must change how we assess students to support our expectations.

Teachers must change how and what they teach and how they assess student progress. New forms of assessment must support changes in classrooms, and the assessments must be tied to the standards. Neither standards nor testing is sufficient alone as reform strategies; they must operate together and be supported by changes in incentives, professional development, technology, early intervention strategies, and reducing health and social barriers to learning.

2. Is the Council advocating a national test and a national curriculum?

There have been conflicting interpretations of the Council's report, and some reflect misunderstandings of the Council's charge and recommendations. For example, at the previous hearing before the subcommittee, Dr. Walt Haney indicated that no one test can accomplish multiple purposes such as system accountability and instructional improvement. In fact, the report specifically states that no one test can serve all purposes and therefore recommends a system of assessments. For system monitoring, NAEP should be changed to reflect the national standards and to include new forms of assessment. For instructional purposes, multiple individual assessments should be developed by a variety of agencies.

3. Why should we have national student standards but not national school or system standards?

The role of the proposed new Council would be national certification of student standards, but the States would have responsibility for delivery standards. National standards are needed for national competitiveness and for equity reasons. How the standards are implemented in schools is a local responsibility because teachers must be involved in decisions about teaching and curriculum. The delivery standards need to originate at the local levels and should be the States' responsibility.

The issue of delivery standards was extensively debated. While national standards and assessments of the standards are the catalysts and are necessary to move our schools toward higher achievement, States and local schools need to be responsible to provide the curriculum and teaching needed by their students. Without involvement and commitment at the classroom and school levels, the standards will not meaningfully impact practice. Further, the standards are to guide instruction, not to provide a national curriculum.

I personally have a very strong concern for equity. I believe that we must provide equal opportunity to learn. We need all three components of reform to achieve high expectations for all students. We need high standards, authentic assessments, and the systemic reform of schools to provide the opportunity to attain high standards. I believe that it is the responsibility of the States to ensure opportunity to learn and develop school delivery standards.

4. Under current conditions, how will we afford to implement national standards and tests?

We are the most tested country in the world. But for all the resources devoted to testing, we have little information that can be used to gauge our progress as individuals or as a Nation. The considerable resources currently devoted to testing and curriculum must be redirected, as well as supplemented.

The current patchwork of tests is not only expensive in terms of financial cost, it is expensive in terms of time taken away from learning in the classroom. The tests are also inefficient, because they provide little useful information.

New forms of assessments, "break-the-mold" or authentic performance assessments, are designed to be used for instruction. They would be more than cost effective and also more efficient.

Professional organizations and the Department of Education are helping to develop the standards, and test development is underway in many States and districts and by commercial publishers as well as research institutions.

I encourage the members to incorporate the recommendations of the Council into legislation to move the Nation toward raising standards for American education.

Chairman KILDEE. I'd like to ask one quick question before we go on to Mr. Geiger, just to pick up on one thing. You mentioned Joe in the fourth grade not doing well on his math test and what should Joe do. Suppose 60 percent of the kids in that fourth grade don't do well on the math test? It's not so much what should Joe do; it's what should the school do; right?

Governor ROMER. Yes.

Chairman KILDEE. So, we really have to look at the school and look at the resources in that school also, then.

Governor ROMER. Right. One of the tests of character of this country is whether or not we've got the willpower to really accept the fact that our performance is as far away as it really is from where it ought to be. Let's assume 60 percent of that class or of the students in the whole system are not doing well, given the standard we arrive at. I think that is truth.

One of the most important things for a free people is to have truth so they can then act upon that truth. When they see that their youngsters are that far away from being competitive in a world marketplace, they then are going to respond.

I believe this country can respond. We simply are living in a kind of a fool's paradise that we think the good life is going to come to us without the effort that has to be put in to get it.

Chairman KILDEE. Knowledge is power, and we shouldn't fear knowledge; right?

Governor ROMER. That's right.

Chairman KILDEE. Mr. Geiger?

Mr. GEIGER. Thank you, Mr. Chairman, members of the committee. I'm Keith Geiger, president of the National Education Association.

I want you to know I'm tempted to discuss this morning the week that I just spent in Japan attending a conference where the Japanese Teachers Union is discussing the need to move from a 6-day student week to a 5-day student week.

I want you to know that come April of this year 240 schools in Japan are moving to a 5-day student week. While we are very curiously watching and looking at them, they are also looking at us because they are very, very concerned about what it is they are doing to their students and what it is they are doing to their teachers in making them come in for six days instead of five.

If I were to make a prediction, I would predict that in the next 5 years you're going to see some things happening in some other countries just like you see it happening here. But that's not what I'm here for this morning.

It is now time for the United States to turn its attention to national security, defined in terms of quality of life not warheads or armed personnel in the field. We must be realistic about the needs of America's public schools. We must let actions speak louder than words.

NEA has long supported efforts to raise standards for educational outcomes to ensure that students, teachers, and schools have clear and high expectations. But until we set standards for the quality of school programs and services, in other words having schools ready for the students also or, as Representative Kildee called it, standards for our schools, for the quality of school programs and services and new measurement of outcomes, these will only confirm what we already know.

Inequalities and inadequacies exist. I concur 100 percent with what Governor Romer said. We've got to do these simultaneously. We can't do one without the other. They have to go hand in hand.

Last session, this subcommittee held hearings on the need for general financial aid to public elementary and secondary schools. We urge this subcommittee to advance legislation that would provide some $100 billion to the Nation's schools.

We believe such funds are necessary to meet the national education goals adopted by the Nation's Governors and endorsed by the president. General aid would enhance the ability of school districts to create break-the-mold schools. It would expand the options available to students and parents. It would help protect programs now being cut as a result of severe economic constraints at the State and local level.

Funds under a general aid to education program would assure that all schools could meet standards for program quality with the following basic elements, and we believe these are basic to any school. America's public schools must have a qualified teacher in every classroom. That means every rural school and every urban school, as well as the suburban schools.

Teachers should be relieved of duties that take them away from time spent teaching students. Third, America's public schools must have reasonable standards for adequate class size, including much smaller class size in the elementary grades.

While I appreciate what Representative Henry said a few minutes ago about not getting one before the other and having options, I'm not sure what options the elementary school teacher in Wales, Massachusetts, who has 75 kindergarten, first, and second graders has or the school that has three teachers with 199 students.

That's a reality in this country today. Kindergarten, first, second grade combined with 75 students, I don't think that teacher has many options on how she teaches.

Public school course offerings should reflect the broad array of skills America needs to maintain our Nation's economic and democratic institutions. I might add extracurricular activities such as art, music and physical education should be an integral part of every child's program. It ought not to be a frill for only a few. It ought to be part of every school program.

Next, every school should have a well stocked library and age-appropriate materials and access to computer-based data systems. I am amazed that we are talking in this country on one hand about better reading for our students, and on the other hand, we are closing our libraries on Saturdays and Sundays so that the students can't visit those libraries.

Next, every school district must have the resources to provide adequate, appropriate facilities and staff with the preparation and

skills to meet the individual needs of students with disabilities. I might add that the Federal share of students with disabilities has fallen from 12 percent to 7 percent in the 1980's.

Next, every school should have properly licensed counselors able to assist with academic, emotional and social issues and health care professionals to address medical emergencies and other health care needs.

Next, every school should set aside the time and resources to provide relevant in-service education for all staff and to work together to define the mission of school and plan its implementation. Teachers want to become better teachers, but it doesn't happen by osmosis.

The math teachers in the Unites States have just established a new set of standards. The math teachers in the United States will not be able to relate to those new standards without intensive in-service training, whether it be on Saturdays, whether it be in the evenings or whether it be in the summer months. They aren't going to learn the new standards without that.

Last, every public school student should be able to attend schools safe from drugs and violence and from environmental hazards. I acknowledge that this index lists the factors that should be included rather than an absolute standard for each element. Developing such a standard with the input of a broad array of stakeholders including teachers, school administrators, State lawmakers and parents would be a useful exercise.

NEA looks forward to working with the subcommittee to refine such an index so that all Americans can know what it takes to look and to demand in our public schools. I might indicate that I also served on the National Council Committee with Governor Romer, with Mr. Kildee, with Mr. Goodling, and I might indicate that what I'm telling you today is nothing that I haven't said at every meeting of that subcommittee that I attended.

I strongly support putting together standards at a national level, putting together an assessment system, not a national test, and Governor Romer clarified that very clearly, but an assessment system which would allow us to see whether or not we are reaching those standards.

I also indicated there, I will indicate here, and I will indicate every place I can that I believe that we have to also set standards, because if Harold Hodgkinson's data is correct, that 33 percent of the youngsters that attend school today at age 5 and into kindergarten are ill-prepared and are doomed to failure before they even enter school, then we have to also have our schools ready for students as well as an assessment process.

Thank you.

[The prepared statement of Keith Geiger follows:]

STATEMENT OF KEITH GEIGER, PRESIDENT, NEA

Mr. Chairman and members of the committee:

I am Keith Geiger, President of the National Education Association which represents more than 2 million professional and support employees in the Nation's public elementary, secondary, vocational, and postsecondary education institutions. This hearing comes at a pivotal juncture in our Nation's history. It is now possible and necessary for the United States to turn its attention from a bipolar international

conflict to our own national security—security defined in terms of quality of life, not warheads or armed personnel in the field.

It is time to be realistic about the needs of America's public schools and to let actions speak louder than words. We support efforts to improve the measurements of educational outcomes: test scores, graduation rates and evaluations of student, teacher, and school performance. But until we set standards for what goes into the school—the quality of the programs schools are able to provide—measurements of outcomes will only confirm the inequalities and inadequacies that presently exist.

As Americans look around them, they see a wide array of long-neglected needs: education, health care, housing, law enforcement, environmental safety, transportation, and more. The ability of our Nation to maintain our place in the international economy is increasingly challenging. And far too many Americans do not have the resources to provide for their families' welfare.

In part, our national optimism has worked against us. The attitude that better times are always just ahead allowed us to borrow against the future to pay for current needs, rather than making an investment in programs that build for tomorrow. We must learn to temper our optimism with a healthy dose of reality. Unless we invest today in programs that strengthen families and communities, we will not maintain our economic strength or national security.

The National Education Association believes that American public schools must play a critical part in rebuilding communities, in renewing a sense of national purpose, in reinvigorating the national economy. And we believe that the role of the Federal Government must be more than coach or cheerleader. America's public schools need more help from the Federal Government than can be provided in a speech.

The education reform movement of the past decade centered on standards. NEA and its affiliates have been active in efforts to raise standards for students, teachers, and schools. But let's face facts. If all students do not have access to quality educational opportunity, then high standards are not a bridge to academic and economic success; they are a barrier.

Last session, the subcommittee held hearings on the need for general financial aid to public elementary and secondary schools. Providing significant resources to local education agencies would go a long way toward enhancing flexibility, creativity, and diversity in schools—while making a major contribution to meeting the National Education Goals adopted by the Nation's governors and endorsed by President Bush.

A vast expansion of Federal aid to public schools is not as radical an idea as some suggest. Some 44 years ago, Ohio Republican Senator Robert Taft explained why he strongly supported Federal aid to education:

"In matters affecting the necessities of life—and I should like to confine it so far as possible to the necessities of life, namely, to relief, to education, to health and to housing—I do not believe the Federal Government can say it has no interest, and say to the people, 'Go your way and do the best you can.'

"Because of the way wealth is distributed in the United States I think we have a responsibility to see if we can eliminate hardship, poverty, and inequality of opportunity, to the best of our ability. I do not believe we are able to do it without a Federal aid system."

The Federal Government made a commitment some 25 years ago to meet the needs of disadvantaged students, provide significant assistance to innovative education programs, and fund research to develop better approaches to teaching and learning. Congress and the White House have never fully met that commitment. In the area of elementary and secondary education alone, the present budget for the U.S. Department of Education is some $15 billion below what it would take to serve all students eligible and in need of federally supported education services.

But meeting those needs alone would not be enough to bring the type of rapid significant progress toward the National Education Goals that the American people want and expect. To provide the quality of services Americans expect and deserve will require some $100 billion in additional Federal funds.

Additional Federal resources—not geared toward specific school populations—should be provided. Chapter I concentration grants provide an excellent illustration of how successful such a general aid program can be. Schools that have received concentration grants have had the flexibility to add programs, staff, and services that help students meet higher academic standards. Such broad-based school improvement programs demonstrate the effectiveness of a concerted school-based reform, as opposed to relying solely on programs targeted to select portions of the school population.

At present, school districts—faced with reductions in State aid and local revenues—are cutting programs, in many cases eliminating those programs we need our

young people to go in most, such as advanced math and science programs and foreign languages.

General aid would enhance the ability of school districts to create "break the mold" schools by having longer days or restructuring the school day, by integrating related topics—such as history and geography—or working with local manufacturers on brainstorming solutions to technological problems.

In addition, general aid would surely expand the options available to students and parents. It would be a welcome change from the choices that currently face local school administrators: do we cut the heat or the staff, do we postpone maintenance on the buses or put off buying textbooks another year?

Without venturing to state what the absolute measure for a school system would be, I would like to offer a quality index that could be used as an outline to define the type of educational environment every student should have:

Teaching staff: *America's public schools must have a qualified teacher in every classroom.*

America's schools will never be able to meet the public's high expectations for academic excellence without strong educational teams. Every school district must have the resources to attract and retain qualified individuals through compensation and benefits commensurate with the preparation and demands of the job. Those who entered the teaching profession 20 or 30 years ago—when job prospects for women were limited—are now leaving. And local school districts are going to have to be able to pay significantly more to attract new, qualified teachers. The teacher shortage problem in this country is manifest in a variety of ways: shortages of teachers in certain geographic areas, particularly in rural and urban communities, and in certain academic disciplines, including math and science, foreign languages, and teaching students with learning and physical disabilities. At the same time, there is a keen shortage of minority teachers; minorities represent about one-third of the school population, but less than one-tenth of the teaching force.

Education Support Employees: *Teachers should be relieved of duties that take them away from time spent teaching students.*

Every school should have the resources to attract and retain a diverse staff, including people properly trained to operate equipment and do other tasks essential to the safety of schoolchildren, and paraprofessional classroom aides to work with children, assist with paperwork, and other tasks as necessary.

Class size: *America's public schools must have reasonable standards for adequate class size, including much smaller class size in the elementary grades.*

There is no one single ideal standard for class size that can be applied from kindergarten through 12th grade. In some instances, it may be desirable to have a larger class size in some secondary programs for a lecture course in concert with smaller sessions for lab work or discussion. But one thing is clear: smaller classes help schools meet many important academic objectives. With small classes, teachers are better able to spend time one-on-one with students, understanding each student's learning style and determining how best to reach them. Smaller classes also allow for greater interaction among students and between teacher and students which leads to developing higher order skills, such as problem-solving and innovation, and social skills necessary to tomorrow's workplace, such as collaboration.

Course diversity: *Public school course offerings should reflect the broad array of skills America needs to maintain our Nation's economic and democratic institutions.*

American public school systems are frequently presented with a difficult choice driven by exigencies of the budget: to cut programs with high enrollment rates, but that are highly challenging, or cut programs that are highly challenging, but with low enrollment rates. Moreover, many schools are forced to cut more and more extracurricular programs. Extracurricular activities should be seen as an integral part of education, an enhancement, not a frill. When American parents talk about choice in education, more often than not they are talking about wanting access for their children to rigorous, quality education programs—not a particular site,

Library: *Every school should have a well-stocked library with age-appropriate materials and access to computer-based data systems.*

Books, magazines, reference materials, and computer-based data systems are essential to processing and synthesizing the complex information available to curious minds. At the same time, school libraries should be a source of materials that let young people read for enjoyment. A well-stocked library truly democratizes education.

Staff and facilities for disabled students: *Every school district must have the resources to provide adequate, appropriate facilities and staff with the preparation and skills to meet the individual needs of students with disabilities.*

Educational programs for individuals with disabilities are a moral and economic imperative. Such programs enhance the self-sufficiency of those served and gives them marketable skills. And yet, the Federal Government has fallen woefully short of honoring the commitment it made to fund 40 percent of the costs of education for individuals with disabilities. In fact, the Federal share fell from 12 percent to 7 percent during the 1980's. Inadequate Federal funding for these programs drains resources for programs that serve all students.

Counselors and Health Care Professionals: *Every school should have properly licensed counselors able to assist with academic, emotional, and social issues and health care professionals to address medical emergencies and other health care needs.*

Many American children face a daunting set of challenges in their personal lives that teachers alone are ill-equipped to address. Many school-age children have parents who are children themselves, parents who have chemical dependency problems, or parents without the resources to provide their basic physical and emotional needs. The schools alone cannot meet all of the basic human needs of children, but neither can school employees ignore them. Public schools can, with properly trained and licensed staff, help coordinate services to students provided by other community agencies.

Instructional Leadership: *Every school should set aside the time and resources to provide relevant inservice education for all staff and to work together to define the mission of the school and plan its implementation.*

Consistent with efforts to create "break the mold" schools, each school should be afforded the opportunity to work in a collaborative fashion on defining the mission of the school and setting forth a plan to meet the needs of the school and the community it serves. In addition, the school's inservice education component should be oriented toward meeting those goals.

Physical plant: *Every public school student should be able to attend schools safe from drugs and violence and from environmental hazards.*

The Education Writers of America's 1989 report, "Wolves at the Schoolhouse Door," detailed the dismal condition of America's public school facilities. EWA determined that one-fourth of all public school buildings need major repair work, and that the cost of new construction—to meet enrollment increases and replace outmoded and, in many instances, unsafe school facilities would cost $84 billion. Many American schools still present hazards to the lives and health of the students and staff from dangerous levels of asbestos, radon, and lead in drinking water.

I acknowledge that this index provides only factors that should be taken into consideration, rather than an absolute standard. Developing such a standard, with the input of a broad array of stakeholders—including teachers, school administrators, State lawmakers, and parents—would be a useful exercise. NEA looks forward to working with the subcommittee to refine such an index with an eye toward letting all Americans know what to look for and what to demand of their public schools.

Thank you.

Chairman KILDEE. Thank you, Mr. Geiger. We will move on to the next witness, Bella Rosenberg.

Ms. ROSENBERG. Thank you. Al Shanker regrets that he couldn't appear before you personally today, and I thank you for letting me represent him for the American Federation of Teachers.

Before I do a summary of my remarks, let me make a couple of observations. In the first instance, and Governor Romer pointed this out, the phrase "national standards and examinations" has become like the word restructuring. It's like a rorschach. People could make it mean whatever they want it to mean. By national standards and examinations, we do not mean a single test at a single point in time that precedes the development of standards.

Another observation that I'd like to make is that we've heard a lot today about how we cannot go about perfecting the world until the world is perfect. We've heard about teacher, Mrs. Simont. We've heard about school financing equities. We've heard about a whole host of things that we have been moaning and groaning about for oh so many years and have barely tinkered around the margins in trying to correct.

These are very, very serious problems. Let me suggest to you that one of the values of the NCEST report is that it suggests a new dynamic for beginning to address those problems. It suggests a discipline for viewing those problems and targeting resources and namely the issue of what do we want our students to know and be able to do and how well are they doing.

Indeed, of all the education reform reports that have been issued over the past several years, few, if any, are more important than the NCEST report precisely because it engages an issue that we have ducked all these years and have essentially ceded to private textbook and testing companies. That is the issue of what do we want our students to know and be able to do.

That attention is long overdue. The public has been out of that discussion, and this hearing, indeed, is a signal that the public is back in that discussion. The NCEST report should not be the end of that discussion. It doesn't answer all the questions. There are lots of problems it doesn't even raise, but it should be the beginning of this most fundamental discussion in education.

Now, you might say we already have content and performance standards, and that's correct. We have high standards; for example, AP requirements and exams for a relative handful and for the vast majority we have the low standards underlying the requirement in most of our States and districts that you have to pass a 6th of 7th grade tests in order to get a high school diploma.

Of course, all of our States and most of our districts have thick books of goals written either in such excruciating detail or vague motherhood-and-apple-pie generalities that no teacher can use them and no parent can understand them, if they've even seen them.

They are, however, very amenable to constructing bureaucratic checklists. So we've conducted an education reform movement for 9 years without ever engaging the most fundamental issue of all, which is what do we want our students to know and be able to do, at what levels of performance, and how will we know if our students are meeting those levels of performance, and how our system as a whole is performing.

There were exceptions to this, of course, during the last couple of years, but so few of them in States and districts that we can, in fact, name them. Most States and districts simply issued volumes of mandates and regulations whose message was "do more and do it better and do it this way."

But more of what and what's better? The answer was always more of what's on the standardized tests or minimum competency tests, and better meant getting up those average scores. It's no wonder that the results of 9 years of education reform have been disappointing to so many people.

Now we have national education goals. We want to meet world class standards but don't know what that means. We want our students to demonstrate competency in challenging subjects, but what content and skills do we want our students to master in the vast fields of English, history, science, and math? What does competency mean? What is good enough, outstanding, needs improvement?

Without answers to these and other questions, we may as well scuttle the national education goals and bury them in the grave-

yard where all other lofty education rhetoric has gone. Therein lies the significance of the NCEST report in our judgment.

For one, it takes the national education goals seriously and by so doing, creates a whole new dynamic for our education system. It also recommends a process for answering these questions in a uniquely American way; that is, without Federal control over education.

It also happens to offer some of the best advice available on a host of issues that have bedeviled education for years: how to join excellence with equity, how to have standards without standardization, how to rationalize the targeting of resources and build the capacity of the system, how to sort out the various uses of testing, uses from abuses, how to discipline the politics of education without sacrificing democratic control, and how to balance professional authority with public involvement.

Now, NCEST does not suggest solutions to all of our problems. It doesn't deal with childhood poverty, one of the most significant educational issues of our day. It doesn't solve the scandal of our school finance system. It also leaves lots of technical problems about standards and assessments still to be sorted out; for example, calibration, equation, all these words that we've been talking about, how you report out incomparable ways from the individual to the national level.

We have some ideas about how to do that, but we don't quite know how to do that yet. It also is a little weak on the system delivery standards, and I won't repeat what Keith Geiger has said. We agree with what he said. But this report is the beginning of the conversation and not the end.

It holds the promise of such a giant step forward and offers such a systemic rather than piecemeal approach that it's hard to believe that more tension is being devoted to everything that could go wrong rather than how to do things right.

Now, what are some of the criticisms we are hearing? Let me just cover a couple of them. One is that in a pluralistic society, no one has the right to say what students should know and be able to do as a result of their education. With that kind of reasoning, we might as well give up public education at taxpayer's expense.

We've also heard that this is going to lead to a national curriculum. It is not going to lead to a national curriculum. There's a world of difference between curriculum and curriculum frameworks. It's also the case that we already have a national curriculum, and it's generally set up by private companies. In fact, we have two curriculum: one for advantaged students and one for disadvantaged students. That's a fundamental inequality that it's about time we started addressing.

We've also heard that it's going to narrow the curriculum and that teachers will only teach those things that might be on the assessments. In the first instance, the curriculum could do with some narrowing. We have so much breath and so little depth and we do need some decision rules because time is finite.

In the second instance, people are afraid that it will drive music and art and all those kinds of subjects out of the schools because there's been no discussion of assessments in these subjects. First of all, we can do something about that. Second of all, in all the other

countries that have these kinds of systems, if anything, they do more music and art and those kind of subjects than we do in ours.

Now, what about local control, our great tradition of local control? In the first instance, it's mostly a fiction. There's very little local control left. Look at where the money for education comes from. Look at the influence of Chapter One testing requirements on local practice. Our tradition of local control doesn't look quite as robust as we like to talk about it.

If anything, national standards can rationalize our policies and target our resources to the local level and reinvigorate local control by putting substance back into district and school decisionmaking. The most serious criticism of all is that national standards in any system of assessment will destroy the progress we've made in equity.

In the first instance, we reject the argument that all poor children are poor achievers, and we reject the argument that poor children can't meet higher standards. We also reject the argument that there will be incentives created to dump poor youngsters into low level tracks or somehow ignore them to concentrate on ways of getting up the school average.

In the first instance, it's a monstrous view of the adults in this system. In the second view, it completely ignores the role of incentives and the control we have over designing incentives to discourage that kind of behavior and punish that kind of behavior where we find it.

That, of course, brings us to the issue of accountability, which is perhaps the biggest worry of the critics of the NCEST report and ours too. There is no question that the nature of the accountability systems we design will determine incentives. Those incentives will shape the behavior both of the adults and youngsters in our education system.

If our accountability systems are rigid, mechanistic and politically driven, which is what they are now, and if new assessments are a substitute for the development of standards, and giving teachers and students opportunities to work with them, then we will likely be no better off than we are now.

By the same token, if there is no accountability based on the standards and assessments we develop, or if there's accountability only for the adults in the system, then we may as well not spend all the time, resources, and energy it will take to properly develop the system.

In our view, the accountability discussion in the NCEST report is perhaps the least satisfying section. It concentrates on the accountability of schools. In effect, by not exposing the fiction of accountability we have now gives a signal that we can overlay that mentality under the new system we develop. That is unacceptable.

The accountability discussions in that report also tend to reify the view that students have no responsibility for their own education. That is, if they do well, it is because the schools did well. If the students fail to achieve, it is because the system failed them.

What we are saying is that assessments have to count for students as well as the adults in the system. We are not talking about tracking youngsters, denying them schooling or job opportunities,

or making go/no-go decisions on their lives on the basis of a single test or a cutoff score.

That has been a gross mischaracterization of what it means to have incentives for youngsters to work hard in school and to make school count. But unless we just like the sound of the phrase that "students are the workers in our schools," and unless we believe that you can get educated without working at it, we had better revisit the issue of student's responsibility for their own learning.

The AFT strongly endorses the recommendations of the NCEST report and urges Congress to act on it. We can either paralyze ourselves into inaction by obsessing on everything that might go wrong and act as if what we have now is good, or we can devote our energies to moving towards a national system of standards assessments in an intelligent, responsible and open and self-correcting fashion.

We urge you to initiate the latter course because inaction means the status quo, and that is simply unacceptable.

Thank you.

[The prepared statement of Al Shanker follows:]

STATEMENT OF ALBERT SHANKER, PRESIDENT, AMERICAN FEDERATION OF TEACHERS

On behalf of the 780,000 members of the American Federation of Teachers, I appreciate the opportunity to submit testimony on the report of the National Council on Education Standards and Testing [NCEST], on which I served, to this distinguished committee.

Of all the umpteen education reform reports that have been issued over the past several years, there are few, if any, more important than the NCEST report. Its subject, content and performance standards and the evaluation of system and student progress toward meeting those standards, is about as basic and essential to teaching and learning and to education policymaking as you can get. And yet we have failed, even refused, to deal with these fundamental issues and have chosen instead to defer to and misuse a substantively impoverished, commercial but cheap, standardized testing system to determine the meaning of education and the nature of accountability for us. My strongest wish is that the NCEST report means that this way of running an education system will soon be over.

The fact that the 32 diverse members of NCEST, and the substantial number of additional people who served on subcommittees, agreed that national education standards and assessments were both desirable and feasible is already a good sign. As recently as a few years ago, anyone who dared to talk publicly about national standards and assessments for American schools risked being considered a radical or a nut. Sure, every one of our competitor nations has such a system, he would have heard. And sure, their students achieve at much higher levels than ours. And sure, every time we try to explain how we deal with curriculum and tests and accountability, foreign visitors scratch their heads in amazement. But, he would have been told, education in those countries is under the control of their central governments while we have local control. Those countries are homogeneous but ours is diverse. Their education systems are elitist and have rigid tracks, while ours is a mass system dedicated to equity. Education in those countries means memorization and recall and working children into premature stress, but we prize creativity, letting children be children, building their self-esteem and enabling them to learn how to learn.

We have repeated these things so often and reflexively that we have failed to notice that they have become mostly fictions. We, too, have central control of education—ceded by the Federal Government, States, and districts to a handful of private test and textbook publishers. And what does local control mean in light of all the mandates and regulations that have rained down on districts in recent years, almost none of which give schools and students a clue about what's expected of them except to beat last year's performance on the standardized test?

We have also failed to notice that most of the nations we're competing against are no longer quite so homogeneous and that diversity has not disabled them from deciding what students should know and be able to do or lessened popular support for standards. Nor are these systems elitist, in the ways we like to think. Japan,

France, Germany or Korea, for example, are as committed to mass education as we are, and they have managed to raise both the floor and the ceiling of student achievement and, in many cases, to get their average student performance to the level of our top performers.

Tracking? Virtually none of our competitors tracks or otherwise labels children in the early grades, which is where we concentrate our tracking practices and begin to send youngsters the message that achievement is a matter of ability—not effort. In fact, we track as much if not more than our competitors (much of it in the later grades by student "choice"), but unlike them fail to give youngsters in the bottom tracks an education of value. And because of our lack of standards, we're not doing too well with the top, either.

Nor does our lack of standards make our children happy in school. In fact, most of our youngsters say that they hate school and are bored and alienated, while children in our competitor nations report they enjoy school, feel engaged and are eager to work hard. Remember, too, that our child suicide rate surpasses that of Japan's, the nation that presumably has the most stressed-out children of all. To be sure, our students lead the world in self-esteem; none of our competitor nations can boast students with such high regard for their own knowledge, skills and achievement. The problem is that our students' self-regard is not matched by their performance. Little wonder, too: We adults in the education system have abdicated our responsibility to students by failing to set standards for their performance.

Now, however, we are beginning to realize the heavy price we have paid for maintaining these fictions about our system and ducking the fundamental issue of standards and the assessments that should support those standards. If we are disappointed with the results of 9 years' worth of education reform activity across the States, then consider that, during that whole time and with few exceptions, not a single State or district ever engaged the issue of what students should know and be able to do, except in the grossest generalities. Instead they issued volumes of regulations whose message was do more and do it better. More of what? What's better? The implicit and, often, explicit answer was do more of what's on the standardized test and make these scores better. And so our teachers did, only to be damned for not having high enough expectations of their students.

If it is hard to imagine how we could have conducted an education reform movement all these years without deciding what we wanted our students to know and be able to do, it is even harder to imagine how we will meet the national education goals related to student achievement without so deciding. "By the year 2000, American students will leave grades four, eight and 12 having demonstrated competency in challenging subject matter including English, mathematics, science, history and geography. . . ." But what content and skills do we want our students to master in these vast fields? What does competency mean, and how will we know it when we see it, let alone communicate to our students what it is? What is good enough, outstanding, unacceptable performance? And what do we mean when we say "the academic performance of elementary and secondary students will increase significantly in every quartile . . ."? Quartiles on which assessments and pegged to what kinds of standards? Without answers to these and other questions, we may as well scuttle the national education goals and consign them to the graveyard in which all other lofty education rhetoric has been buried.

And therein lies the significance of the NCEST report: It takes the national education goals seriously, and it suggests a process for answering all these questions in a uniquely American way that does not involve Federal control over education. Indeed, in my judgment the NCEST report offers the best advice currently available on a host of questions that have bedeviled our education system for years: How to join excellence with equity; how to have standards without standardization; how to build the capacity of the system and rationalize the targeting of resources; how to sort out testing for accountability purposes from testing for instructional purposes; and how to discipline the politics of education without sacrificing democratic control.

Of course, the NCEST report does not suggest a solution for all of our problems. It does not tell us how to overcome childhood poverty, one of the great drags on our students' achievement. And it leaves lots of difficult issues, like comparability of exams and reporting comparable results from the individual to the national levels, among others, still to be sorted out.

But the NCEST report is such a giant step forward that it's puzzling to me that more attention has been devoted to what could go wrong with a national system of standards and assessments than to what we need to do to make sure such a system is properly implemented and monitored.

What are some of the charges we have heard?

The first is that, in a pluralistic society, no one has the right to say what all students should know and be able to do or to set standards for performance. With that line of reasoning, we may as well give up running an education system at taxpayers' expense.

The next thing we hear is that national standards will lead to a national curriculum and narrow teaching and learning. The fact is we already have a national curriculum—the low-level one that's set by a handful of testing and textbook companies. Indeed, we have two national curricula: a steady and boring diet of basic skills for disadvantaged students and a richer fare for advantaged students. That's a fundamental inequality, and it's unacceptable. As for national standards and assessments narrowing the curriculum, I'm all for that. Our curriculum is very wide indeed, with breadth of coverage substituting for depth. The only legitimate aspect of the narrowing curriculum argument that I can see is that schools will only teach those subjects for which there are standards and assessments. That is indeed undesirable; it is also a problem that is easily headed off, as the experience of other nations and a little common sense tell us.

What about local control? There's precious little of it already, and look where its gotten us: unequal access to high-quality education. If anything, national standards, and rationalizing our policies and targeting of resources on the basis of those standards and local and school progress toward meeting them, will reinvigorate local control and put substance back into district and school decision-making.

Nor need we worry about stifling teacher creativity. We have been systematically driving out creativity and undermining professionalism for years by telling teachers exactly what they must do virtually every minute of the day and subjecting them to every passing educational fad. Having teachers participate in developing national standards and the curricula based on those standards is the first step back towards creativity, not away from it. Letting teachers know what we expect students to know and be able to do, orienting teacher education and professional development toward those standards, and then letting teachers work together to figure out what methods work best for their students are more likely to unleash creativity than continuing to hold them hostage to standardized tests.

Neither can I accept the argument that national standards and assessments will further disadvantage poor children, especially in light of the NCEST report's strong words on the equity issue. I do not accept the implicit argument that poor youngsters are all low achievers or that poor children are unable to meet higher standards; I do not accept that schools will dump low-achieving youngsters into dead-end tracks on the basis of new assessments or ignore them in order to concentrate on getting better-performing youngsters to do even better so that a school's average score looks good. That is a monstrous view of the adults in our system. And that will only happen if we design incentives to encourage and reward that behavior instead of discouraging and punishing it.

And that of course brings us to the issue of accountability, perhaps the biggest worry of the critics of the NCEST report and mine, too. There is absolutely no question that the nature of the accountability systems we design will determine incentives and those incentives will shape the behavior both of the adults and youngsters in our education system. If our accountability systems are rigid, mechanistic and politically-driven and if new assessments are a substitute for the development of standards and giving teachers and students opportunities to work with them, then we will likely be no better off than we are now.

By the same token, if there is no accountability based on the standards and assessments we develop or if there is accountability only for the adults in the system, then we may as well not spend all the time, resources and energy it will take to properly develop this system.

In my view, the accountability discussion in the NCEST report is perhaps the least satisfying section. It concentrates on the accountability of schools and, by not exposing the idiocy of most current accountability schemes, does not discourage the overlay of our present, educationally bankrupt accountability mentality onto a new system. Much more work needs to be done on this issue.

The accountability discussions in the report also tend to reify the view that students have no responsibility for their own education—that if they do well, it is because the schools did well; and if the students fail to achieve, it is because the system failed them. What I am saying is that assessments have to count for students, as well as the adults in the system. I am not talking about tracking youngsters, denying them schooling or job opportunities or making go/no go decisions on their lives on the basis of a single test or cutoff score. That has been a gross mischaracterization of my position and of what it means to have incentives for youngsters to work hard in school and to make school count. But unless we just like the

sound of the phrase that students are the workers in our schools and unless we believe that you can get educated without working at it, we had better revisit the issue of students' responsibility for their own learning.

I strongly endorse the recommendations of the NCEST report and urge Congress to act on it. We can either paralyze ourselves into inaction by obsessing on everything that might go wrong or we can devote our energies to moving towards a national system of standards and assessments in an intelligent and responsible and open fashion. The AFT chooses the latter course; the status quo, which is what inaction means, is unacceptable.

Chairman KILDEE. Thank you.

Dr. Koretz?

Mr. KORETZ. Thank you, Mr. Chairman, and thank you for the opportunity to be here. Before I begin, I'd like to make two quick comments. First, I'd like to ask that my complete statement be submitted for the record. I'd also like to point out that I'm representing, as one of your colleagues put it, the unified team.

I prepared this statement jointly with three other people: George Madaus and Albert Beaton of the Center for the Study of Testing Evaluation and Educational Policy of Boston College; and Edward Haertel of Stanford University. All three are nationally renowned scholars of educational measurement. I'm privileged to be presenting on their behalf.

We share some of the premises of the NCEST report. We agree that standards are too low in many American schools, and we think the NCEST report is to be commended for giving the debate about educational standards the kind of political prominence it hasn't had in a very long time.

But nonetheless, we are very deeply troubled by the report and by the policies it recommends. Even though the policies it recommends seem common-sensical to many people, we believe that they are unlikely to work and that they may well have very serious side effects as well.

We all know that the problems in American schools are complex in their causes and have proved very difficult in the past to ameliorate. But the Council's report, when you come down to the essence, is very simple: standards linked to high stakes tests with serious consequences for individuals, especially kids.

It does make passing reference, and only passing reference, to other things that are needed such as improved professional development and the reduction of health and social barriers to learning, but it's proposals are tests and standards, and it leaves the rest to States and unspecified others.

This is not a new idea. The idea of driving education with accountability going to tests has been tried for many centuries in many countries, and its track record is generally quite poor. I won't dwell on this at great lengths. Several people have already mentioned some of this.

But holding people accountable for performance on tests tends to narrow the curriculum. It tends to inflate test scores, leading to phony accountability. It can have pernicious affects on instruction. It has, in some cases, had adverse affects on students already at risk, such as increasing the dropout rate and leading to greater cramming in high minority schools than in others.

The question is, as the Council said, we must take care to avoid unintended and undesired affects of some testing practices such as

these, but how? The Council suggests four ways. We're skeptical of all four. I'll treat them in turn.

By analogy, it's very much as if the Council came to you and said well, we don't like the fact that some of our kids are sick. We want your support to medicate nine million of them a year. It is true that the medicine we propose to you bears a kind of uncanny resemblance to ones that have failed in the past. No, we haven't finished developing the new one. We haven't even begun testing it, but let's go ahead anyway.

In no area outside of education would we even seriously consider such a proposal. The first of the four ways that NCEST proposes something different that will avoid, in theory, the negative affects of test-based accountability that we've seen so often before is the idea of setting up standards that will only be a core. They won't be a national curriculum. They will be a core that teachers and administrators elsewhere will elaborate upon.

Well, the evidence suggests otherwise, and the question is why should we expect it to be different this time? The historical evidence is very clear, as is the research evidence. When you hold people accountable for performance on a test, they spend more time testing what's on it.

The other side of that is they spend less time testing what's not on it. What's not on it, we know from research, is often quite important. In other words, it narrows the curriculum. Now, perhaps Ms. Rosenberg is not bothered by that, but I am and my colleagues are because one of the consequences of that is inflated test scores.

In other words, kids do better on tests than they actually do in terms of what they learn. The public is mislead as a result. The evidence shows that this kind of inflation is widespread, and our own research shows that it can be egregious.

This kind of inflation of test scores should have been an essential concern of the NCEST report but was not for two reasons. First of all, it undermines the validity of the tests. If tests are telling you that kids know more than they do, ipso facto they're not valid.

Moreover, it makes a mockery of the Council's call for greater accountability. Certainly, teachers and students are held accountable but policymakers are not. They're let off the hook because test scores make the public believe that things are better than they actually are. We've seen that wholesale in the United States in the 1980's.

Well, the NCEST report also suggests a second way to deviate from current performance in the hopes of doing better. It's to rely on new types of tests, performance assessments. Governor Romer gave a very articulate and impassioned defense of performance assessments.

I would like to start by saying that the four of us don't oppose the development of those new types of assessments; in fact, several of us are actively involved in doing so. But it's necessary to temper our enthusiasm with a bit of realism.

First, what NCEST proposes is, in part, a return to the past. For centuries, testing programs have been evolving toward greater standardization, not away from it, for very simple reasons: cost, practicality, a desire for comparability, and a desire for objectivity.

They are good reasons to criticize what we're currently doing. They are certainly good reasons to criticize the rampant misuse of standardized tests that we see in the United States but we shouldn't be blind to the fact that if we start soft-peddling goals such as objectivity, there will be a very serious price to be paid, and there are some difficult problems to address. The Council report does not seriously address them.

Second, to the extent that these new performance assessments really are innovative, most of them are incompletely developed at this point, and most of them are entirely untested. They are ready for R&D. They are not ready to be the linchpin of a national policy. We don't believe in flying an airplane while it's being built, particularly when the passengers are children who have no choice about being on board.

Third, the field actually knows something about performance assessments. Not all of them are new. What we know suggests we ought to be cautious. I'll just summarize this very briefly because we've all been here a long time now. I'll just give you three points and I can return to these in questions, if you'd like.

First, student performance is often quite inconsistent from one performance assessment to another, even when they measure in theory the same thing. What that means is they tend to give us in many cases unreliable estimates of what students know and estimates that are often of questionable validity.

Second, performance assessments are no more immune to the problems of teaching to the test than any other kind of test. It may be that we can devise tests that are so interesting that preparing kids for them will be good instruction. I'll leave that question aside. But nonetheless, if you teach to any test, you run the risk of inflating the scores, regardless of whether the test is multiple choice or performance based.

There's also no evidence, as far as we know, that switching to performance assessments will lessen inequities in the current testing program. Finally, the proposed reliance on performance assessment raises very difficult questions of feasibility that the Council reports simply didn't address.

For instance, recent experience in Great Britain, which has been experimenting with a system that bears some striking resemblances as to what NCEST asks for showed a need for extra staff and support in schools, a serious disruption of classroom activities, difficulties in rating large numbers of performance tests, and difficulties in establishing comparable enough conditions for testing across schools so that scores actually mean the same thing.

Third, the NCEST report calls for a system of examinations. I guess I have to dwell on that since that's become a bit of a theme today: a system of exams rather than a single national exam. That system of exams has an extraordinarily wide range of functions.

It's supposed to be bottoms up. It's top down only in the sense that it's linked somehow to standards. We think the Council's expectations for this system are unrealistic and in some cases mutually contradictory. I can only discuss a few of the problems here today.

First, and perhaps foremost, given the discussion that's already happened, we don't believe that the level of diversity that the

Council has discussed in its deliberations will provide any protection from the negative affects of externally imposed high stakes exams. It avoids the political albatross of a national test but not its practical consequences.

If you were to look at Europe as an example, which the Council has often suggested doing, European exam systems, in fact, involve considerable diversity. France and Germany, if they were a model, would suggest up to 40 examining boards just for secondary students in the United States. We're not convinced that even that would in fact give any effective voice in curriculum to teachers, building principals, or even, in many cases, local administrators. It's voluntary, but one would ask for whom.

The system is also supposed to provide, despite the diversity of tests that it produces, comparability of scores. The Council report is adamant on that point. It says that comparability of results is essential. We think that this is one of the most puzzling and glaring holes in the Council's report. Tests are supposed to be the same but different. We, frankly, don't know how that can be done.

The Council doesn't say just how similar tests should be, but it does note, for instance, that they want these tests to be used to screen kids for employment or screen kids for admission to postsecondary education. That means it ought to be very similar indeed, because if they are not, they will be inequitable.

Students who happen to land a bad test will be capriciously denied employment. At the same time, these tests that, by that standard, have to be the same, have to be different because they have to be linked to local curricula which would evolve independently.

The Council's way out of this seeming contradiction seems to be that this test will be the same in some core respects linked to the standards but not in others. We consider that an intriguing idea, but at this point, it really doesn't have any flesh on its bones. We think it would be prudent to wait to find out what it means before we make it a cornerstone of national policy.

Finally, research gives us a very clear reason to be cautious about this notion of different tests being the same. That is, even fairly minor differences among tests can produce dramatic differences in the results that they provide. The differences are often matters of equity.

For example, changing tests in even fairly moderate ways can change the apparent difference between blacks and whites or males and females. In our view, that's not a trivial concern.

Since we're now to the topic of equity, I'll be very brief about this because it has been discussed on both sides quite a bit already. But we're very concerned about inequities in the proposed system. First, we agree with the Council that students cannot be fairly held to the same standard if they're not given equal opportunities to learn.

Unfortunately, the Council report doesn't flesh out what delivery standards are and it doesn't set up any mechanism for getting them on the table quickly.

Second, as the Council's own implementation task force noted in very explicit detail, many low achieving students face obstacles outside of school. The Council's implementation task force used very

strong language. For example, saying that it's essential. The words were "the health system must be restructured to better meet the needs of the poor."

Somehow that got lost before the final report because the Council report doesn't suggest any national actions to do anything about those inequities.

Finally, to go back to your comment, Mr. Kildee, we were also very puzzled by the fact that the Council report says it's all right to impose national standards to hold kids and teachers accountable but not to hold States accountable. In our view, that's a reversal of the traditional Federal role which is to establish or guarantee equity of services for disadvantaged groups.

We think that if States cannot be trusted to set standards for kids, perhaps they shouldn't be trusted to set them for the delivery of services either.

Finally, costs. The Council report says that detailed cost estimates are unavailable. It says, as Governor Romer repeated today, that the new assessment shouldn't really add to the net burden of teaching. We find that, frankly, hard to swallow.

There are data on costs of different kinds of assessment, and they suggest that the new system would be very expensive indeed. For example, I have several here, but I'll just give you one in the interest of brevity. Let's take the college board AP exams, which the Council report itself uses as an example.

Those currently cost $65 a kid per subject. For the five subjects recommended by the Council, that's $325 a kid, which compares to—Michael Kean could correct me if I'm wrong—I believe about $2 to $5 a kid for current testing. Well, testing in three grades, which is what the Council recommends, would run at that price over $3 billion a year, which is not chump change.

But that's not the worst of it, because if tests are really going to be used to influence instruction, not just to sort out low ability kids, which is what they're used for in many of our industrial competitors, it may not be enough to test in one grade out of four.

AP exams certainly aren't used that way. Every year-long course is tested at the end of that year. So, we may well have to test, if we're going to follow that model, more than three times in a kid's career, maybe six, seven, eight times. Now, we're getting up to $6, $7, $8 billion.

That's not the only basis for expecting high costs. European exam systems also do, as does their own national assessment. Our two questions are this: Is that kind of money going to be available; and second, even more important, if it is, where is the discussion of whether testing is the best use of that money at a time when school districts and State education agencies are going bust.

The Council report simply doesn't say. Finally, in passing, let me just say that experience elsewhere shows that the nonfinancial costs of the kinds of assessments that the Council recommends are going to be sizable, costs like professional development, disruption of classrooms, foregone instructional time, other allocations of teacher time, and so on.

One last comment on the Council report, and then I'd like to give you some alternatives, is the proposal which is now an S2 to establish something called a National Education Standards and Assess-

ment Council to provide quality assurance. This is the fourth major way that the Council report would deviate from current practice.

NESAC, if you will pardon the acronym, is supposed to certify tests and assessments. We believe that as currently structured in the NCEST report and in S2, the certification of assessments provided by NESAC would be a simple sham.

Discharging the responsibilities that NESAC has given would require substantial, substantive, and technical expertise, but the recommendations do not call for the appointments of so much as one person who has that expertise. It doesn't provide NESAC-needed independence, and moreover the recommendations show a thorough-going misunderstanding of what is needed to validate and evaluate tests.

It can't be done merely by having a group of informed people sitting around the table looking at tests. It has to be done by looking at the impact of those tests in the actual context in which they are used. That takes time. It takes serious empirical work. It takes serious research. The NESAC model does not provide anything, any provisions, for commissioning, funding, or using the kinds of research that are needed.

Well, what are some alternative directions? Despite our criticisms of the report, we do believe that standards and assessment will play a very important role in national education reform, and that they should. We do not——

Let me refer back to the hearing you had on February 4th where one of the Council members accused people who opposed the NCEST report of "thinking that it's better to stick with the discriminatory and educationally destructive current testing technology rather than invest in developing the new technical capacity we will need," which was echoed to some degree by Ms. Rosenberg.

Count us out. All of us have been strong critics of the present system. The question is, how to do better. I'll give you five concrete suggestions which I'd be happy to elaborate upon. First, establish standards and do it now. But establishing standards has to go well beyond the kind of general prescriptions that are in something like the National Council of Teachers of Mathematics' standards.

They must be specific enough, in other words, they must be curricula, to guide assessment and teaching. Those have to be the first steps because you can't fairly hold students accountable for learning something if you can't demonstrate that they've been taught it.

Second, support research and development. A serious R&D effort is exactly what we need now to answer the questions of desirability, feasibility, validity, fairness, and so on that the Council report, in our view, didn't answer.

That R&D effort has to go well beyond simply designing interesting tests. It has to include, for example, conducting its serious empirical research on the quality of these new tests, conducting investigations of their costs, both indirect and direct, conducting research into the effects of these new types of assessment affects on instruction, on learning, on school organization, and on equity, and finally, building an infrastructure capable of supporting this whole edifice.

Third, learn from small-scale implementations. Governor Romer has repeatedly expressed concern that States are getting ahead of

us. We consider that a gift. We ought to be encouraging local and State innovations with the proviso that unlike most of them that are going on now, they'd be seriously evaluated so that at the end of that time, we will actually have some information that we can use to fashion a national system if we want one. We believe that approach can be easily accommodated within the framework of H.R. 3320.

Fourth, specify, implement, and document the other components of reform. I'll be very brief here, but Governor Romer has said that tests and standards are only the bread of the sandwich, not worth much without the filling. We absolutely agree, but we think it's time to start talking about what that filling is and creating it.

We need clarification of what these school delivery standards are and we need to set up the mechanisms that will start providing equitable services for kids.

Finally, we need to establish realistic workable procedures for evaluating this whole thing. We need an independent body with sufficient expertise and credibility to evaluate the quality of assessments, to examine their consequences, and to judge the comparability of their results.

To do that, a group would have to have substantial technical and substantive expertise. It would have to be independent. It would have to have the authority and the funding to commission extra research. Its charter would have to call for a serious evaluation of the effects of assessment programs on schooling, learning, and on various groups of children, especially disadvantaged children.

Regardless of that, regardless if anything is done with NESAC or a successor to it, there is no reason not to build—and, in fact, there are many reasons to build—realistic evaluation into what we're doing already. We recommend that reforms funded by H.R. 3320 or other similar vehicles be required to do the following, if they use assessment as a tool of reform: to specify in advance precisely how tests will be used; to specify what evidence of reliability and validity will be collected, if they intend to use innovative assessments; and to require that grantees planning to use test-based accountability in any form whatever be required to evaluate its affects on instruction, its affects on diverse groups of students, and the possible inflation of test scores.

To conclude, I'm sorry I've gone on longer than I should have, when all of that is done, we will have developed standards in curricula. We will have produced sorely needed information about the affects of the new types of assessments that the Council has asked for. We will have put into place a mechanism that can actually help protect against both inadvertent harm and outright abuse.

At that time, we'll be able to make rational decisions about a national system. We can't know what those decisions should be now because the evidence isn't all in, but we suspect, based on all of the extensive evidence already in hand, that the answer will be, in part, that assessment should be a partner in reform but can't be its primary engine.

Mr. Chairman, thank you for the opportunity to address you today.

[The prepared statement of Daniel M. Koretz, Ph.D. follows:]

STATEMENT OF

Daniel M. Koretz
Senior Social Scientist
RAND

George F. Madaus
Boisi Professor of Education
Director, Center for the Study of Testing, Evaluation, and Educational Policy
Boston College

Edward Haertel
Associate Professor, School of Education
Stanford University

Albert E. Beaton
Professor of Education and Senior Research Fellow,
Center for the Study of Testing, Evaluation, and Educational Policy
Boston College

before the

Subcommittee on Elementary, Secondary,
and Vocational Education
Committee on Education and Labor
U.S. House of Representatives

February 19, 1992

Preparation of this testimony was supported by the RAND Institute on Education and Training and The Center for the Study of Testing, Evaluation, and Educational Policy at Boston College. The opinions expressed here, however, are solely those of the authors.

Mr. Chairman and members of the Subcommittee, thank you for the opportunity to speak with you about the recent report of the National Council on Education Standards and Testing (NCEST) and about the critical issues of national standards and tests.

Today I will present a statement that I have prepared jointly with Dr. George Madaus and Dr. Albert Beaton, both of the Center for the Study of Testing, Evaluation, and Educational Policy at Boston College, and Dr. Edward Haertel of Stanford University. Dr. Madaus and I both served on NCEST's Assessment Task Force. Drs. Madaus, Beaton, and Haertel are nationally recognized scholars of educational measurement, and I am privileged to present this joint statement on their behalf. Drs. Madaus and Beaton are here today; Dr. Haertel was unable to attend.

We share some of the premises and goals of the National Council, and we commend the report for giving them a rare prominence in political debate. Standards *are* too low in many American schools. We share the Council's concern that expectations have been especially low for groups that have historically done poorly in school. A nationwide debate on educational standards, carried out properly, could have substantial positive effects on education and should begin promptly.

Nonetheless, we are deeply troubled by the NCEST report and the policies it recommends. Although the NCEST recommendations may appear commonsensical, they are unlikely to work, and they may well have serious negative side effects. Moreover, in our view, the NCEST report does not adequately address key issues of feasibility, fairness, validity, and reliability--precisely the issues emphasized in the Congressional charge to the Council.

I will use the limited time today to discuss only a few critical aspects of the NCEST recommendations, and thereafter I will discuss very briefly a few alternatives. We believe that standards and assessment *can* play a key role in educational reform, but to fulfill that potential and avoid negative side-effects, they must be used in ways quite different from NCEST's recommendations.

A HISTORICAL PERSPECTIVE ON THE NCEST RECOMMENDATIONS

We know from many years of research that the problems of American education have many causes and are often very difficult to ameliorate. Yet NCEST's recommendation is far simpler than these difficult problems warrant: standards linked to "high-stakes" tests that have serious consequences for individuals. True, the report notes that tests and standards

are not "panaceas" and lists in passing some other needed elements of reform, such as improved professional development and "the reduction of health and social barriers to learning" (p. 7). But it puts its faith in tests, maintaining that tests and standards "can be the cornerstone of the fundamental, systemic reform necessary to reform schools" (p. 5). Moreover, its specific proposals for national action are largely limited to tests and standards. Most of the remaining and more difficult aspects of school improvement, such as professional development and family and community supports (p. 37), are left for states and localities, and the report offers no specific proposals for dealing with them.

Using test-based accountability to drive education is hardly a new idea. This approach has been tried many times over a period of centuries in numerous countries, and its track record is unimpressive. Most recently, it was the linchpin of the educational reform movement of the 1980s, the failure of which provides much of the impetus for the current wave of reform, including the Council's report. Holding people accountable for performance on tests tends to narrow the curriculum. It inflates test scores, leading to phony accountability. It can have pernicious effects on instruction, such as substitution of cramming for teaching. Evidence also indicates that it can adversely affect students already at risk--for example, increasing the dropout rate and producing more egregious cramming for the tests in schools with large minority enrollments. NCEST has proposed some departures from current testing practice in the hopes of doing better, but those departures are unproven and hold as much peril as promise.

Despite its Congressional charge, the Council report did not discuss the evidence about test-based accountability, but it acknowledges past problems fleetingly in asserting that "care must be taken to avoid the unintended and undesired effects of some testing practices, such as narrowing instruction and excluding certain students from instruction" (p. 29). Indeed, but how? The burden rests with the Council to explain why we should have confidence that this hoary prescription will work so much better this time. In our view, the Council has not made a persuasive case.

By analogy, it is as though the Council came to you and said: we want your support to medicate 9 million children a year because we don't like their being ill, but the medicine we propose using bears an unsettling similarity to some that have failed and had serious side effects in the past, and we have not yet finished designing the new medicine or testing it. In no field other than education would we consider, let alone accept, such a proposal.

HOW WOULD THE NCEST APPROACH DIFFER FROM PAST PROGRAMS?

The NCEST report proposes a system of testing that differs in several respects from current practice. Specifically:

- The standards would not be a curriculum that would control what people teach; rather, they would be a "core" upon which educators would elaborate.
- The proposed system would rely substantially on different types of tests: "performance assessments."
- NCEST is proposing a *system* of tests that would be under local, state, or "cluster" control and would be free to differ but that would nonetheless be linked to national standards.
- A new entity, the National Education Standards and Assessment Council (NESAC), would provide "quality assurance."

The Council maintains that these four attributes will make the old prescription of test-based accountability work better than previously. We are skeptical of all four.

WILL THE NEW STANDARDS BE JUST A "CORE?" THE RISK OF NARROWED INSTRUCTION

Advocates of the new testing often say that "what you test is what you get." Their first premise is that when tests are made to matter, people teach what you test. Their second premise is that we can get them to teach better by testing better things and giving the tests serious consequences. This is the core logic of the NCEST report.

We agree with the first of these premises. The historical and research evidence, both in the United States and elsewhere, consistently shows that when people are held accountable for performance on tests, teachers focus on the tests' content. The problem is that to spend more time teaching what *is* tested, teachers generally spend *less* time teaching what is *not* tested, and what they give up is often important.

That is, external tests coupled with serious consequences generally narrow the curriculum. The NCEST report deals with this problem only obliquely. It maintains that the new *standards* will be only a "common core" that would be enhanced by local elaboration. Perhaps, but the essential problem lies not with the standards but with the tests that would be used to

implement them, and the Council report does not explain why these tests will not narrow the curriculum.

One consequence of narrowed instruction is the *inflation of test scores.* That is, students do much better on a particular test than their actual mastery of the subject matter warrants, so the public is misled about students' and schools' performance. This inflation of scores appears widespread, and our own research indicates that it can be egregious.

Inflation of test scores should have been a central concern of the NCEST report for two reasons: it undermines the validity of the tests, and it makes a mockery of the Council's goal of greater accountability. First, if a test suggests that students know more than they actually do, it is *ipso facto* invalid. Second, when scores are inflated on a high-stakes test, students and teachers are indeed held accountable, but policymakers, such as chief state school officers and governors, are let off the hook because performance appears better than it really is.

Despite the critical importance of these issues, the NCEST report had little to say about them, apart from holding out a vague promise of new types of tests. To what extent can current technologies help counter the problem? How should inflation of scores be detected? What should NESAC do about the inflation of scores when deciding whether to certify a test? The Council report is mute.

WILL NEW TYPES OF ASSESSMENTS SOLVE THE PROBLEM?

The NCEST report pins its hopes on the proposed use of "innovative" types of performance-based assessments. These tests, we are told, will focus on "higher order or complex thinking skills" (p. 28) and will be "worth teaching to" (p. 6).

We do not oppose the development of innovative forms of assessment; indeed, some of us have been involved in that effort. It is essential, however, to temper our enthusiasm with a bit of realism.

First, what NCEST proposes is in some respects innovation, but in other ways it is a return to the past. Over the centuries, testing programs have evolved and become more standardized for reasons of cost, practicality, administrative convenience, and a desire for comparability and objectivity. There are ample grounds for criticizing current objective tests--and in particular the misuse of such tests that recently has become commonplace. But to change direction and to downplay concerns such as objectivity and comparability for the sake of other goals--say, richness of assessment or

better incentives for teachers--will confront us with very serious difficulties that the Council report alludes to but does not adequately address.

Second, to the extent that the proposed assessments really *are* innovative, they are in many cases unfinished and untested. They are at a stage where they are ripe for a serious R&D effort, complete with rigorous evaluation, but they are not yet ready to be a linchpin of national policy. There are many practical, technical, and infrastructure issues that must be resolved before such techniques can safely be deployed as policy instruments on a large scale in schools. We do not believe in flying an airplane while building it, particularly when the passengers are children.

Third, the field already has considerable knowledge of how certain performance assessments have worked, and the evidence to date suggests the need for caution.

To put the evidence about performance assessments into perspective, we need to recall what tests are. Regardless of the types of exercises it comprises--multiple-choice questions, essays, experiments, performance tasks, or whatever--a test is only a *sample* of student knowledge. It is useful *only if we can generalize from performance on the test to a broader domain of interest*, such as "mastery of algebra" or "understanding the nature and process of science." One key to reliable and valid generalization is having an adequately large and representative sample of exercises.

If we cannot draw valid generalizations from the exercises on the test, two things follow: the public is misled about student performance, and--if the test has real consequences--people are treated capriciously. That is, some students who fail on the basis of one overly limited or non-representative sample of tasks would have passed if given an equally defensible alternative set. In other words, both *validity* and *fairness* are undermined.

Evidence suggests that student performance often generalizes poorly across related performance tasks. The quality of student essays, for example, varies markedly depending on the type of essay required and even the specific prompt used. Similar findings have come from recent investigations of hands-on science tests. In the case of multiple-choice and short-answer tests, it is relatively easy to deal with problems such as this by adding additional questions. Performance tasks, however, tend to be costly to produce and time-consuming to administer and score, so using a large set of exercises is less practical.

One partial solution to this problem, if we are willing to pay the large development costs, would be to use a large set of performance tasks across a large group of students--say, all those in a state or large district--but to

administer only one or a few to each child. This approach, however, generally does *not* give us valid and reliable scores for individual children. The NCEST report calls for comparable tests at all levels down to the individual student and for serious consequences for individual students, but it leaves us in the dark about how this is to be accomplished with fairness, validity, reliability, and at reasonable cost using complex performances that show limited generalizability.

The validity of performance assessments can also be undermined by teaching to the test. NCEST argues that the new tests will be designed to be taught to. By this, proponents usually mean that teaching to these new tests will itself be good instruction. Perhaps, but that is only one of the reasons to be worried about teaching to the test. The other, as I have already noted, is that teaching to the test narrows instruction, thus inflating scores and undermining validity. To our knowledge, there is no evidence that performance tests are less susceptible to this problem than conventional tests, and there are some indications that they are more susceptible.

Switching to new forms of assessment is also unlikely to help make the system more equitable. Some proponents argue that switching to performance assessments will lessen inequities in the current testing system. Again we know of no evidence substantiating that claim, and some evidence suggests the reverse.

Finally, the proposed reliance on new performance assessments raises serious questions of feasibility that the Council did not address. Great Britain's recent experiment with Standard Assessment Tasks (SATs) provides an illuminating example. Among the specific issues that arose in England and Wales were the need for extra support and staff in schools, the need for procedures to minimize the disruption of school and classroom organization, the difficulty (and perhaps undesirability) of imposing standardized conditions of administration that would permit comparability of results across schools, and the difficulties inherent in rating large numbers of performance-based tasks.

CAN WE BUILD A SYSTEM OF ASSESSMENTS THAT CAN DO WHAT NCEST PROPOSES?

The NCEST proposal calls for a *system* of examinations that will serve an extraordinarily wide range of functions. Among them are providing comparable information across jurisdictions and at all levels of aggregation; providing incentives to educators to teach better and to students to work harder; and providing valid predictive information to select students for further education or employment. Moreover, this system of tests would be "bottoms-up;" only the standards to which they are somehow linked would be

"top down." Many of the Council's expectations for this system are unrealistic, and some are mutually contradictory. Moreover, the Council's list of functions focuses primarily on secondary school students, leaving unaddressed the difficult question of the uses and consequences of the tests the Council recommends administering in the elementary grades.

Providing Diversity and Protecting Local Initiative

We believe that the degree of diversification needed to preserve local initiative is far greater than the Council envisioned. Absent that diversity, the system will act as a national examination and will not be able to avoid the pernicious consequences of large-scale, external, high-stakes examination systems. European systems have often been suggested as an exemplar, but they entail more diversity than is commonly recognized. In the former West Germany, for example, with a population roughly comparable to that of California, New York, Ohio and Missouri, there are eleven state Ministries for Education that set the separate Abiturs for each state. In France, with a population about the size of that in California, New York, and Ohio, there are 23 separate academies that set the Brevet de college exams and the Baccalaureat exams for each academy. For the United States, a comparable level of diversification would imply the creation of between 18 and 40 examining boards for the secondary level alone. And with over 15,000 school districts, it is questionable whether even this level of differentiation would give teachers, schools, or communities an effective voice in curriculum.

Providing Comparability of Scores

Permitting a great deal of diversification would also run counter to the Council's assertion that *comparability* of results is essential. This is one of the most glaring and puzzling holes in the Council's position. The new tests are supposed to be different but yet the same. How different, and in what sense comparable? The Council never clarified explicitly what it means by "comparable," but it appears to want a rigorous standard of similarity. After all, if tests are going to be used to determine high school graduation, admission to college, or employment (p. 5), they had better measure very similar knowledge and skills. Otherwise, the system will be inherently inequitable, capriciously favoring students taking one exam over those taking another.

Yet at the same time, the new system is supposed to protect local autonomy and encourage development of diverse local curricula. This implies that the tests linked to these curricula would have to be quite different. The Council report seems to be saying that the way out of this apparent contradiction is that the tests and curricula will be the same in some "core" respects having to do with national standards but will be free to

differ in other respects. This notion is intriguing but still lacks any real substance, and it would be prudent to wait to see what this actually means before making it a cornerstone of national policy.

Research gives us even more reason to be cautious: even fairly minor differences between tests can produce fundamental differences in their results. These differences can be a matter of equity; for example, changing tests can alter the apparent size of differences among racial and ethnic groups or between males and females. Moreover, teaching to the test--an explicit goal of the Council report--can greatly exacerbate differences in performance from one test to another.

Equity

We have serious concerns about inequities in the proposed system. NCEST recognized two potential sources of inequity, but we are not satisfied by its responses to either.

First, students cannot be held fairly to the same standard of performance if they are given unequal opportunities to learn. The NCEST report acknowledges this in its call for "school delivery standards," but it would let others figure out what these ought to be, and it says nothing about the resources that would be needed to attain reasonable equity.

Second, many low-achieving students face barriers to achievement that lie outside of school, and the NCEST proposal does not seriously address them. For example, as the Council's own Implementation Task Force noted, "Students whose most elemental shelter, food, and nurturing needs are not met are not going to perform at the minimum competency level consistently, much less demonstrate...world class academic standards. The health...system...*must* be restructured to better meet the needs of the poor..." (p. G-11, emphasis added). That message appears to have been lost between the Task Force report and the final NCEST report. Many statistics show how poorly this nation addresses those elemental needs--e.g., figures on our appalling infant mortality rates, our high rate of child-poverty, the large numbers of children without access even to basic health and dental care, and so on. Yet the Council report proposes nothing specific to alleviate those inequities.

NCEST provided no persuasive rationale for using a national mechanism for holding students and teachers accountable but not for holding states and localities accountable for providing them with equitable support or educational services. Why can we trust the states to worry about delivery standards but not student performance standards? This would reverse the traditional division of responsibilities, in which the federal government's

interventions in education have often been designed to create equity in the delivery of services.

Costs

The likely costs of the proposed system should also give us pause. The Council says that detailed cost estimates are unavailable but that the new assessment system should not add to the net burden of testing (p. 31). There *are* bases for estimating both financial and other costs of the new system, however, and it is readily apparent that the proposed system would add a very large burden indeed.

One basis for a rough cost estimate is the College Board Advanced Placement (AP) Examinations, often cited as an exemplar and noted in the Council report. AP exams currently cost $65 per subject, or $325 per student for the five-subject battery proposed by NCEST. (By contrast, a commercial standardized test battery costs about $2 to $5 per student.) Assuming that economies of scale are offset by the cost of increased reliance on performance tasks, this suggests a cost of more than $3 billion per year for testing only in the three grades suggested by NCEST.

But that is not the worst of it. It may not be practical to limit exams to three grades if they are to guide instruction (rather than simply weeding out less able students, which is the primary function of exams in some other countries). Certainly an AP-style exam system would require more frequent testing; AP exams are tightly linked to one-year course syllabi. So, let's say we test in six grades instead of three; that raises the ante to over $6 billion per year.

The AP exams are not the only basis for expecting high costs. Recent experience in Europe suggests costs of about $135 per student just for scoring 4 to 5 essay exams, each comprising 4 to 6 questions, when the exams are graded by teachers and include no performance tasks. Our own National Assessment, which requires a mere hour per student and has traditionally been mostly multiple-choice and machine-scorable, has cost roughly $100 per student.

Who is going to provide these billions of dollars? Perhaps even more important, if that amount of money were made available, would examinations be the most effective way of spending it? Neither of these issues is addressed by the Council report.

The non-financial costs of the proposed system are likely to be substantial as well. One cost will be forgone instructional time. For example, in Great Britain, the recent administration of SATs was supposed

to extend over three weeks, but some local education agencies maintained that it would take them more than six weeks. In other parts of Europe, preparation for exams stretches over a period of three months to a semester. Advocates of the new testing assert that the new exams would be so challenging that preparation and testing time would be good instruction, not time taken away from instruction. Long experience to the contrary suggests that we should wait for substantiation before accepting this assertion, especially given that our nation already has an unusually short instructional year. Experience also suggests that innovative examination systems will require time-consuming and expensive inservice training in addition to substantial teacher time for preparation, administration, and logistics.

Gauging the Effectiveness of Schools

Finally, the new system is supposed to be able to "provide evidence about the success of schools, local school systems, states, and the Nation in bringing all students...to high performance standards" (p. 13). Here we need to draw a distinction that the Council did not. It is entirely feasible to build a system that will monitor the progress of students at various levels of educational organization. The National Assessment does that for the nation as a whole and is experimenting with doing so for states, and there is no reason why that system could not be expanded and enriched.

Such data, however, generally tell us *little about the effectiveness of schools or systems.* They tell us which groups are doing better, but not why. Current assessment systems, and the new system proposed by the Council, simply provide the wrong sort of data to evaluate programs. To evaluate educational programs, one must be able to rule out plausible alternative explanations of performance differences. This is the same, simple standard used in all of empirical science. To do this with educational programs requires collecting extensive, high-quality data on factors (such as family background) that exert powerful influences on achievement, tracking the movement of students among educational systems, and, in most cases, tracking changes in student performance over time. None of this can be done by testing students in one grade out of four and aggregating to the district or state level. Senator Moynihan's recent tongue-in-cheek argument that states' scores are caused by their distance from the Canadian border was a humorous but powerful reminder of the riskiness of ignoring these simple rules of scientific inference.

WILL NESAC PROVIDE QUALITY ASSURANCE?

The Council report and now S. 2 would give NESAC major responsibilities for quality assurance, such as establishing guidelines for developing assessments, gauging their validity, and ensuring their

comparability. The report calls for NESAC to guide the certification of standards and assessments.

If NESAC is established as proposed by the NCEST and S. 2, however, its certification of assessments would be a sham. Discharging its responsibilities would require substantial substantive and technical expertise, but the recommendations do not call for the appointment of even a single individual with expertise in measurement or evaluation. NESAC would also lack needed independence; its members would be appointed by the new National Education Goals Panel, and certifications would be made jointly by both organizations.

Equally important, the NCEST recommendations show a thorough misunderstanding of what is needed to validate tests and monitor their effects. A test cannot be validated by asking a group of individuals to examine its content, as the NCEST report implies. Moreover, validation is an ongoing process, not a one-time effort. To validate a test requires substantial empirical research, and the NESAC model does not make provisions for commissioning, funding, or using the needed investigations.

To take one example, suppose that a test is used to screen individuals for employment. Then it is essential--for reasons of law and equity, apart from simple ethics--to demonstrate that performance on the test predicts performance on the job. This requires empirical research. The need for such research, although long established in law as well as in the measurement profession, is ignored in the NCEST recommendations.

To take one other example that I alluded to earlier: if a test is used for accountability or is for other reasons "taught to," as the Council report explicitly recommends, how does one know whether scores on the test have been inflated enough to undermine their validity? Having a committee such as NESAC examine the test to see whether it lives up to "world class standards" will not provide a clue. Again one needs research, for example, random substitution of similar tests in ongoing testing programs. Once again, there is no provision in the NCEST recommendations or S. 2 for that type of validation.

WHAT ARE SOME ALTERNATIVE DIRECTIONS?

Despite our criticisms of the assessment system proposed by NCEST, we do believe that standards and assessments shou'd play a role in education reform. Contrary to the accusation leveled at opponents of the NCEST report by one Council member before this Subcommittee, we do *not* believe that "it is better to stick with the discriminatory and educationally destructive current testing technology rather than invest in developing the new technical

capacity we will need for the program outlined [by NCEST]"[1] In fact, all of us have been strong critics of the present system. The question is how to do better.

Establish Standards and Curricula

We endorse the proposal to move ahead with a national debate on educational standards. This effort must go beyond generally worded standards to include the development of curricula specific enough to guide teaching and assessment. *These must be the first steps*; a syllabus-based examination system will have to wait until standards are established, because we cannot insure that students have a fair chance to learn what is tested until we have curricula in place.

There is more to establishing standards, however, than the NCEST proposal envisions. If we want standards that reflect, for example, skills and knowledge that are needed for certain types of jobs or for certain types of postsecondary education, we will need to validate the standards and confirm empirically that the standards actually reflect what is needed.

Support Research, Development, and the Building of Infrastructure

A serious R&D effort is precisely what is needed if we are to answer the questions of desirability, feasibility, validity, practicality, fiscal costs, opportunity costs, educational costs, and consequences that are raised by the NCEST proposals but not seriously addressed in the NCEST report.

This R&D effort must go far beyond the design of new assessments, and it will take considerable resources and time. We must, for example:

- Conduct serious empirical research on the quality of new performance assessments--for example, the reliability of scores and the extent to which they generalize enough to be meaningful.
- Conduct investigations of costs, including non-financial and indirect costs.
- Conduct research into the *effects* of new types of assessments--effects on quality of instruction, learning, school organization, and equity.

---

[1] Statement of Lauren B. Resnick before the Subcommittee on Elementary, Secondary, and Vocational Education, Committee on Education and Labor, February 4, 1992.

- o Build an infrastructure capable of supporting new assessment systems.

The R&D effort would need to focus on the *context and use* of the new assessments. Assessments are not good or bad in the abstract; their quality depends on how they are used. A test may succeed in providing one type of information and fail utterly to provide another; it may be beneficial if used in one way and pernicious if used differently.

Learn from Smaller-Scale Implementations

During the NCEST deliberations, Governor Romer repeatedly expressed concern that states and localities are getting ahead of the national effort. State and local efforts should be seen as a gift, not a cause for anxiety. If states and localities can be encouraged to couple their assessment innovations with serious evaluations--as Vermont is now doing--their efforts will provide an invaluable source of information about what works and what doesn't, and this information can improve national efforts. Indeed, the national program should include active encouragement of diverse smaller-scale efforts, but these efforts, unlike most of the recent innovative assessment efforts with which we are familiar, must be coupled with adequate evaluation. We believe that this approach can be accommodated easily within the framework of H.R. 3320.

Specify, Implement, and Document Other Components of Reform

The NCEST report says that tests and standards are not a panacea, and Governor Romer has often noted that they are only the bread of the sandwich, worth very little without the filling. We concur, and we believe that it is time to work on the filling. First, we need further clarification of "school delivery standards:" what specifically must schools provide before we are willing to say that opportunities are equitable? Second, we need to specify *what the other components of reform will be and how they will be implemented*. It is in our view simply unacceptable to hold students accountable for their performance without providing them the opportunities they need to succeed on our examinations. Third, we need to develop and evaluate the indicators of equal opportunity that the NCEST proposal presumes. We know that simple, conventional measures such as teachers' years of experience will not suffice.

Establish Workable Procedures for Evaluation

As I have already noted, we believe that the proposed NESAC would not be capable of evaluating the new standards and examinations meaningfully. We see the need for an independent, non-partisan body with

sufficient expertise and credibility to evaluate the technical qualities of alternative assessments, examine the evidence about their feasibility and costs, monitor the consequences of their use, and judge the comparability of results across the various local and state components of the assessment system. To be effective, such a body would need to differ from the proposed NESAC in many ways:

- Its members must have the needed technical and substantive expertise in measurement, evaluation, and education.
- It must be independent.
- It must have the authority and funding to commission extramural research as needed.
- Its charter must call for realistic validation and evaluation of examinations *as they are actually used in specific contexts.*
- Its charter must call for evaluation of the effects of assessment programs on schooling and learning.
- Its charter must call for evaluation of the effects of the programs on diverse groups of students, particularly the disadvantaged.
- Its charter must call for ongoing evaluations of the strengths and weaknesses of assessment programs, rather than unrealistic, one-time, up-or-down "certification" decisions.

Regardless of decisions about NESAC, evaluation and validation efforts should be built into federally supported education reforms that use assessment. During the 1980s, very few jurisdictions using test-based accountability evaluated the effects of their programs, and some flatly refused outside evaluations. We should not allow this to be repeated. People who want to experiment with dangerous medicines should be required to evaluate the impact of their experiments. We suggest that all reform efforts funded by H.R. 3320 or other similar legislation be required to do the following if assessment is to be a part of the reform:

- Require grantees to specify how tests will be used--in particular, how, if at all, they will be used for accountability at any level of aggregation (from students to states or clusters). Accountability need not entail concrete sanctions; it can be sufficient to publicize scores as an index of performance.

- Require that grantees specify what evidence of reliability and validity will be collected if innovative assessments will be used.

- Require grantees to specify what steps, if any, they will take to lessen the risks of inflated test scores and narrowed instruction.

- Require that grantees planning to use test-based accountability evaluate its effects on instruction, its effects on diverse groups of students (in particular, the disadvantaged), and the possible inflation of test scores.

Because many school districts will lack the expertise to structure reasonable evaluations, it may be helpful to encourage partnerships, for example, between school districts and universities, for these purposes.

Decide About a National System

When all of the steps above have been taken, we will have developed standards and curricula; we will have produced sorely needed information about the feasibility, validity, fairness, and desirability of various types of new assessment programs; and we will have put into place a mechanism that can help protect against abuse as well as inadvertent harm. At that time, the nation will be far better prepared to make reasoned decisions about a possible national assessment system such as the one proposed by NCEST.

We cannot know now precisely what directions will be suggested by information that we do not yet have. The extensive evidence already in hand, however, suggests that a key part of the answer will be to use assessment as a partner in reform, not as its primary engine.

Mr. Chairman, thank you again for the opportunity to discuss these issues with the Subcommittee. At this time, I would be pleased to answer questions.

Chairman KILDEE. Thank you very much.

Dr. Kean?

Mr. KEAN. Chairman Kildee, Mr. Goodling, and members of the subcommittee, I'm very pleased to be here this morning on behalf of the Association of American Publishers Test Committee, the AAP, to address the report of the National Council on Education Standards and Testing.

The AAP applauds the Council's call for high voluntary national education standards and for raising the expectations for our education system. However, we want to insure that a real world workable approach is put in place, an approach that builds on the best available assessment technology and does not reinvent the wheel.

After I describe the areas of the report that concern the AAP, I will offer six specific steps that we suggest that the committee might take in adopting legislation in response to the Council report. First, however, our concerns.

Number one, tests will not revolutionize the education system. AAP must part company with the NCEST report in its focus on break-the-mold assessments. Can any test actually revolutionize the education system? Though test results can be used to target instruction, the test does not increase achievement; good teaching does.

Professional educators know that assessments are built only after specific purpose and specific objectives to be measured have first been identified. Any other way places the cart of testing before the horse of learning. Thinking that new tests will somehow drive educational reform is no wiser than the hope that a new kind of thermometer will lower a fever or cure the common cold.

Number two, a workable sequence of improvement is needed. Publishers strongly disagree with those who say that standard setting and assessment development can occur simultaneously. It is necessary to establish a sequence of events for both professional and for equity reasons.

You can generate all the content standards you want, but if they are not adopted around the country, if new textbooks and instructional materials do not incorporate them, if teachers are not trained to teach them, and finally, if students are not taught the information in them, then any assessment tool based on such standards will fail to accurately measure progress toward them.

Indeed, they will simply measure failure, the failure of teachers, the failure of schools, and the failure of the standards themselves.

The third area of concern, the cost effectiveness of allocatable resources for the development of assessments. It would be a colossal waste of money to build assessments simultaneously with new standards. At the right time, tests built to those standards will be able to do their job assisting and assessing and evaluating student performance and providing teachers with valuable information.

Until then, however, Congress would be throwing good money away on the unsupported assumption that different assessments alone will motivate students and teachers and drive change. The Council failed to assess the cost effectiveness of new assessments.

It presently costs, as Dr. Koretz indicated, a few dollars per student to administer a complete battery of tests. The kinds of tests envisioned by the Council would cost a great deal more.

Fourth, there's no clear primary purpose for the new assessments. The report identifies five generic and a dozen more specific purposes, apparently to be filled by the new assessments, including, as has been mentioned before, controversial high stakes purposes. No assessment can do all things.

Effective assessments are built only after their specific purposes have been identified, not the other way around. Further, what valid purpose is served by imposing high stakes onto assessments traditionally used in elementary and secondary education?

These assessments are achievement tests not SATs or minimum competency tests or IQ tests. They should not be comingled with other, as yet, unidentified purposes, especially if such purposes are linked to high stakes decisions.

A fifth concern relates to the emphasis upon performance assessments versus a multiple measures approach. While the NCEST report recognizes the need for multiple measures of assessment, it expresses unwarranted preferences for performance-based assessments.

Now, less my concerns about this bias be misinterpreted, I should point out that many commercial publishers, my own company included, are and have been for some time very involved in developing and marketing various types of performance-based tests.

Mr. Goodling mentioned a news clipping this morning in Maryland, and I need only point to the State of Maryland, some 10 miles or less down the road, Maryland has one of the largest and one of the most comprehensive State testing programs in the country, making use of multiple measures.

They rely most heavily on new performance assessments comingled with multiple choice tests. Their testing program is also developed by the State of Maryland in consonance with a commercial publisher.

Another concern is with the technical quality of the assessments recommended. Many of the performance-based assessments envisioned in the NCEST report are wholly unproven to replace the range of valid assessments that are used today.

Available research strongly indicates that performance-based assessments do not meet high technical standards. Some have argued that the performance tests do not have to meet these technical standards or that they should be excused from meeting them in order to make better tests available.

I ask you, should we take two steps back in order to make one step forward? All assessments developed to the new national standards must meet the same technical standards, including validity, reliability and fairness. To allow, much less to encourage, the use of any new assessment without these safeguards would be detrimental to the health of our children's education.

Furthermore, performance-based tests also raise serious equity and serious subjectivity concerns. Recent research, including studies presented to Congress by its own Office of Technology Assessment, show problems of racial and gender bias associated with performance-based assessments.

A final concern relates to local options. The NCEST report recommends a centralized top-down governed instructor to certify standards and assessment. Such an approach will delay local op-

tions and the implementation of bottoms-up reform. It should be rejected, especially the notion that specific local or State assessments must be certified first at the national level.

States already require test developers to conform to stringent technical quality standards and to demonstrate tight curriculum alignment. These same competitive processes will work and should be encouraged in developing and in selecting assessments linked to the new national standards.

Now, based upon the concerns that I've just articulated, I'd suggest that the subcommittee consider the following six recommendations: develop standards, determine the assessments purposes, resolve school delivery standards questions, conduct research and alternative assessments, report the results to Congress, and then take appropriate action.

First, develop world class standards. By encouraging the continued development of demanding content in student performance standards, Congress would lay the foundation for improvement of student learning. However, any governing structure approved for this task need provide only a coordinating and an oversight function and should rely on appropriate professional organizations with specific curricular expertise.

Step two, determine the purpose of assessments. Since the Council report fails to focus on the primary purpose for new assessments, the governing body must first deal with this issue head on.

Third is simultaneous resolution of school delivery standards. Once content standards are developed, it will be necessary to implement then. Monitoring implementation requires obtaining information on how many States and school districts have adopted the standards, what types of instructional materials have been produced that incorporate these new standards, what percentage of teachers have been trained in the new standards, how many students are being taught using the new standards, as well as other information on resource availability or distribution, such as the number of computers in schools or the amount of teacher training in assessment methods, et cetera.

Recommendation number four, conduct research on alternative assessment. While publishers, assessment centers and various consortia around the country are already working on a variety of different assessment development projects, more needs to be done.

Congress should encourage additional research on ongoing assessment programs including research to determine how to design new forms of assessment to meet technical standards.

Next, report to Congress. Congress should establish an appropriate time frame within which the governance body is to report on an interim and on a final basis. These dates should be set at the earliest reasonable period, given the importance and given the complexity of the tasks to be performed.

The last step in AAP's recommended process can occur once all of the information has been collected, information on standards and their implementation, on where resources have been committed to enable students to achieve the new standards, and on the standards of assessment research.

When Congress has received and analyzed all of this information, then it will be ready to act on the recommendations about the new

assessments. This approach will avoid the problems and shortcomings we have identified with the NCEST report.

Unfortunately, there is no way to shortcut this process without jeopardizing the very school children who are the subject of the exercise. Some have suggested that a response to the NCEST report should be considered as part of H.R. 3320. However, since the NCEST report is fundamentally inconsistent with the bottom-up approach of the existing bill, we oppose making major adjustments to H.R. 3320.

I believe that these alternative recommendations will create a more effective and a more confluent effort toward the development of national education standards, their implementation throughout the Nation's schools, and the development of assessments linked to those high standards.

Thank you, and I'll be happy to attempt to answer any questions that you might have at this time.

[The prepared statement of Michael H. Kean, Ph.D. follows:]

TESTIMONY OF

DR. MICHAEL H. KEAN

Testimony of Dr. Michael H. Kean before the Subcommittee on Elementary, Secondary and Vocational Education on the Report of the National Council on Education Standards and Testing.

Chairman Kildee and Members of the Subcommittee:

I am Michael Kean, Vice President, Public and Governmental Affairs for CTB Macmillan/McGraw-Hill, one of the country's largest commercial test publishers. For the last seven years, I also have served as chairperson of the Test Committee of the Association of American Publishers ("AAP"), which represents the interests of commercial and non-profit publishers involved in various forms of assessment. I have been personally active in research and development of student assessments, and my background includes ten years as an administrator in a large, urban school district.

I am very pleased to be here this morning on behalf of the AAP to address the Report of the National Council on Education Standards and Testing, entitled Raising Standards for American Education. Although there are some aspects of the Council's Report that test publishers find compelling, other findings and recommendations reflect a less-than-accurate representation of the real world of assessing our students. We want to ensure that a real world, workable approach is put in place, one that builds on the best available assessment technology, provides the critical measurement needed to assess progress on national goals, and does not "reinvent the wheel" when it comes to committing the scarce education resources available today.

After I describe the areas of the Report that concern AAP, I will offer six steps AAP urges this Subcommittee to take should it adopt legislation dealing with the Council Report. I believe that our alternative recommendations will create a more effective and confluent effort toward the development of national education standards, their implementation throughout the nation's schools -- eventually, leading to the development of assessments linked to those high standards.

Identity and Interests of AAP

Mr. Chairman, the AAP represents over 235 publishers of hardback and paperback books, professional, technical, and scientific journals; computer software; and educational and classroom materials, including numerous forms of tests, evaluation scoring materials and instructional materials. As such, AAP members have traditionally been involved in meeting the needs of school districts and state education programs for both standardized multiple-choice and performance-based assessment instruments that have been used to assist in the measurement and evaluation of the nation's school children. While the NCEST report credits publishers for what it terms "promising new efforts," in fact, test publishers have responded to each and every new educational reform movement of the last three decades. Congress should know that AAP members will continue to devote their resources to building tests

BEST COPY AVAILABLE

that reflect new "world-class" standards -- but it is imperative that the process to be followed is appropriate and technically sound.

We should not attempt to "fly the plane" while it is still being built and before it has been fully and thoroughly shown to work. Thus, a structured, cautious approach represents the best opportunity to advance our education system and to improve what students learn in school.

NO CHANGE IN H.R. 3320 IS NEEDED

H.R. 3320 provides the framework for states and local education agencies to carry out the objectives of the NCEST report -- without having to get bogged down in the complicated and unnecessarily-centralized bureaucracy called for in the report. Under H.R. 3320, a state would first adopt goals and curriculum frameworks, and put in place the delivery system to carry them out. Then, the state (and local education agencies) would adopt an assessment program, in which it would identify the purposes to be achieved by their assessments; develop or select the assessments appropriate for those purposes; assure that these assessments reflect the adopted curriculum; have in place a mechanism to assure that they are valid, reliable, and fair for their intended use; and have a monitoring system to assure that the assessments are, in fact, used appropriately and in accordance with the state and local assessment program.

The process envisioned in H.R. 3320 would provide the means by which states and local education agencies would assure that teachers and administrators are trained in test administration and interpretation, as well as in communication of test information to students, parents, and the public -- regardless of the test format.

H.R. 3320 thus preserves state and local primacy, plus maintains a system that encourages flexibility, innovation, and competition. It would not lock states and local education agencies into what ever may be determined to be the conventional wisdom, fad, or panacea of the day -- as determined by a centralized, national arbiter of assessment "correctness." After all, it is within the current system that states and local education agencies have been able to experiment with the alternative assessment formats, including portfolios, demonstrations, and performance. How would any such innovation be accepted, let alone accommodated, under the centralized system proposed by NCEST, now apparently codified in S. 2?

How does one break into the marketplace with a new idea if it does not follow the directives that have been deemed "correct" by a national board, barren of expertise, whose imprimatur -- the official seal of approval -- is necessary in order for that assessment to be used by a state or school system.

Some have suggested that a response to the NCEST report should be considered as part of H.R. 3320, the Neighborhood Schools Improvement Act, already reported favorably by this Subcommittee and the full Committee. We do not support that approach. AAP believes that, while the substance of the NCEST report can be accommodated within the language of H.R. 3320, the governance structure proposed by the Council is fundamentally inconsistent with the "bottom-up" approach advocated in the bill. Therefore, we would oppose making major adjustments to H.R. 3320 in order to make it compatible with the Council's governance recommendations.

NO SIMULTANEOUS STANDARDS AND ASSESSMENTS

AAP applauds the Council's calls for high education standards, particularly in the five key subject areas, and for raising expectations for our education system. AAP also supports related Council recommendations that:

- o voluntary, national and dynamic standards should be adopted; and

- o state and local control over education reform, including the selection of assessments, should continue.

Unfortunately, this second point is largely undermined by the Council's recommended "top-down" governance structure. Significantly, though, only if these conditions are met can the desirability and feasibility of a national approach to setting subject area content standards be undertaken. Accordingly, any legislation this Subcommittee considers should reject a national curriculum and endorse the widest possible participation from individuals and organizations at the state and local levels.

Where AAP must part company with the NCEST report, however, is over its misguided focus on "break-the-mold" assessments (whatever that term means). Can any test actually revolutionize the educational system? Experience has shown the answer is, "No." Though test results can be used to target instruction, the test does not increase achievement -- good teaching does! Those of us who are professional educators know that assessments are built only after a specific purpose and specific objectives to be measured have been identified. Any other way places the cart of testing before the horse of learning. Thinking that new tests will somehow drive educational reform is no wiser than the hope that a new kind of thermometer will lower a fever or cure a cold.

Furthermore, publishers strongly disagree with those who say that simultaneous standards-setting and assessment development should occur. It is necessary to establish a sequence of events because of both professional and equity reasons. You can generate the best content standards possible, but if they are not adopted around the country, if new textbooks and instructional materials

do not incorporate them, if teachers are not trained to teach them, and finally, if students are not taught the information in them, then any assessment tool developed based on them will fail to measure success under those standards. Indeed, assessments built to new "world-class" standards, without some delivery system already in place, are doomed to measure failure -- failure of teachers, schools, and standards.

It might strike some Members of this Subcommittee as odd for test publishers to sit here and tell you not to pass legislation requiring more testing, at least not until certain preliminary steps are taken. Really, it is not all that surprising. Publishers employ hundreds of professional educators and others trained in psychometrics and we know it would be a colossal waste of money to build assessments simultaneously with new standards. Proper sequencing requires that **limited education resources must first be spent on direct student services, including teacher training and instructional materials**. At the right time, tests built to those standards will be able to do their job -- assisting in assessing and evaluating student performance and giving teachers valuable information about what they need to do to improve student learning.

Further, the Council failed to assess the cost-effectiveness of new assessments, one of the four categories it was directed by Congress to consider. It presently costs a few dollars per student to administer a complete battery of standardized tests. The kinds of tests envisioned by the Council would cost at least ten to fifty times as much. In the face of severe national economic and budget constraints, it would be far more appropriate to use the funds we have for textbooks, teachers and other tools that help our children learn.

By accepting simultaneous development of standards and assessments, Congress would be throwing good money away on an unsupported -- and invalid -- assumption, that different assessments alone will motivate students and teachers and drive change. Indeed, we fear that simultaneous development could have serious counterproductive results -- negative effects on learning and impeding education reform at the state and local level.

INTERMEDIATE STEPS FOR ASSESSMENTS

Lest publishers be accused of merely seeking to delay change or prevent improvement, let me clearly state that while content and student performance standards are being developed and school delivery systems are being implemented, areas of assessment work can go forward. Indeed, publishers have already begun to release new tests incorporating the National Council of Teachers of Mathematics standards so heavily praised by the Council. I expect that the same will be true as additional high content standards are developed.

Nevertheless, it also could be very important and worthwhile for additional objective and scientific research and other related activity to go on in the assessment area, even while "world-class" standards are being developed and implemented. However, since AAP has serious problems with several aspects of the Council report as to what needs to be done with assessments, I will detail the three most crucial concerns related to what the focus of such research should be.

1. **There is no clear primary purpose for assessments.**

The NCEST report does not identify clear purposes for which assessments linked to national content standards will be developed and utilized, nor who should make such determinations. In fact, while the report identifies five generic purposes, it actually names twelve specific purposes apparently to be filled by new assessments -- including controversial high stakes purposes. **AAP's caution, however, is that no assessment can do all things.**

This caution, along with the view that purpose should be determined well in advance, is consistent with what the advice given the Council by its Assessment Task Force. Unfortunately, these fundamental considerations were given short shrift by NCEST and were virtually ignored in its final Report.

In addition, AAP has grave reservations about whether any valid purpose is served by imposing high stakes onto the kind of assessments traditionally used in elementary and secondary education. I should point out that our tests are not SAT's or minimum competency tests or IQ tests, but are achievement tests for elementary and high school students. Because these tests are quite different -- and are designed for very different purposes -- Congress should be very concerned about the detriments of commingling test purposes, especially if one of them is high stakes.

An assessment that is valid, reliable and fair safely can satisfy no more than one, possibly two, purposes at once. Moreover, as I indicated earlier, an effective assessment is built only after the specific purpose has been identified, not the other way around. Following the Council's recommendations in this area would result in violating the axiom that no one should be evaluated before knowing what is expected. Tests themselves will not correct problems with what students know.

2. **There is an unwarranted bias in favor of performance-based assessments, without holding them to the same technical standards.**

While the NCEST report recognizes the importance and appropriateness of **multiple measures of assessment**; it expresses unwarranted preferences for performance-based assessments to the

apparent exclusion of other formats. Lest my concerns about this bias in the report be misinterpreted as fear of that form of testing, nothing could be farther from the truth. Publishers (my company included) are very involved in developing and marketing various types of performance-based tests, in addition to the standardized multiple-choice tests that have received so much unfair criticism.

However, publishers have developed performance-based assessments using high quality control to ensure compliance with technical standards. Moreover, publishers evaluate their tests, along with their customers, to ensure that they do not have negative effects on learning.

No achievement test should be administered to students -- whether in a multiple-choice or performance format -- without a prior careful and objective evaluation of its purpose for its effects on learning, its fairness, and its validity and reliability. Certainly no single, unproven assessment format should become the keystone for a national assessment system. If that assessment system is intended to revolutionize education, as well as monitor success in attaining educational standards, and if high stakes are piled on top, then this "brave new world" must be approached with the utmost caution.

Congress should adopt a more logical and deliberate legislative solution and, in the process, refuse to rush the development of so-called "break-the-mold" performance-based assessments. First, many of the performance-based assessments envisioned in the NCEST report are wholly unproven to "replace" the range of proven assessments used today. Indeed, many of these "authentic" assessments are more a part of classroom instruction than they are tests.

Available research strongly indicates that performance-based assessments do not meet high technical standards. Yet, I was amazed during the Council's discussions to hear some professionals argue that performance tests do not have to meet those technical standards -- or that they should be excused from meeting them in order to make "better" tests available. That logic escapes me completely -- and it runs counter to the statements by the American Educational Research Association that its standards already apply to all performance-based tests.

Proponents of performance-based testing, then, cannot have it both ways. Either those tests are really used solely for classroom instruction, in which case they would not have to meet the AERA technical standards -- or they are full-blown assessments used for scientific assessment and comparison purposes and therefore must meet high technical standards.

Any new assessment linked to national content standards, regardless of whether the format is performance-based or multiple-choice, must not be put in place until it meets a thorough review of technical quality. Each assessment must meet high technical quality standards including the hallmark standards of **validity and reliability**. To allow, much less encourage, the use of any new assessment without these safeguards would be detrimental to the quality of our children's education.

3. **Performance-based tests also raise serious questions of equity and subjectivity**.

As some of you are doubtless aware, multiple-choice assessments were developed more than 50 years ago in response to criticism that performance-based forms of measurement in use at the time, were known to be highly subjective -- and discriminatory on both a racial and gender basis. Again, recent research shows that these problems have not disappeared, but actually increase when schools use performance-based tests.

Contrary to the repeated criticisms of current standardized tests, it is completely false to assert that multiple-choice tests can only assess rote learning and memorization. In fact, publishers have been building questions designed to assess how students arrive at and communicate answers and to engage in higher-order thinking skills and problem-solving. To those who ought to know better, I say they should actually read current tests before they dismiss their technical capability.

NO CERTIFICATION OF STANDARDS AND ASSESSMENTS

The NCEST report recommends a centralized,"top-down" governance structure to certify standards and assessments. AAP believes such an approach should be rejected, particularly any notion that specific assessments should be certified at the national level.

It will be very costly to move away from the current practice, which is already a very rigorous approval process. States have been requiring test developers to conform to content and technical quality requirements, as well as to demonstrate curriculum alignment and other information called for in extensive RFP's. These **same competitive processes will work, and should be encouraged, in developing and selecting assessments linked to national standards.** The Council's recommendation will serve only to undermine local options and delay implementation of the "bottom-up" reform already underway at the state and local levels.

In addition to the very detailed and demanding RFP procurement process, our tests are subject to professional standards and independent external review. I emphasize the word "our" because not all tests being administered to students now, or proposed to

be administered in the future, are subject to the same rigorous review. H.R. 3320 is a more appropriate means for assuring that all assessments used for purposes beyond the classroom meet the highest professional and technical standards.

Finally, there is no need to "reinvent the wheel" when it comes to the development of assessments. Private sector test publishers have the resources and expertise necessary to assist state and local development of valid, reliable and fair assessments linked to national content standards. Consequently, Congress should not offer special subsidies or incentives to compete directly with the very responsive, competitive private sector.

A new national regulatory body simply is not necessary, particularly if there are no assurances that the standards developed and used by the proposed Council will be done fairly. There is no guarantee that we will all be playing on a level field.

AAP RECOMMENDATIONS

I have distilled down the serious concerns over the NCEST report into a series of six recommendations for this Subcommittee. These alternatives would enable Congress to take positive, constructive steps in response to the report without fear that it would be inconsistent with H.R. 3320. This would also avoid "unintended consequences" by attempting to build standards and assessments simultaneously -- only to find that there are negative effects on learning or that competition among test developers has been stifled.

Moreover, adopting these steps would provide necessary structure to the amorphous and nebulous concepts laid out by the Council. AAP believes that any new governing body should not be allowed to "define" its own functions.

Step 1: Develop "World-Class" Standards

By encouraging the continuation of development of high content and student performance standards, more or less in accordance with the Council recommendations, Congress would lay a strong foundation or cornerstone for improvement of student learning. AAP would urge, however, that the governance structure approved for this task (and others listed in subsequent steps) need provide only a purely coordinating and oversight function to ensure that standards-setting professional organizations are guided and directed as they follow an approach modelled on the one used by the National Council of Teachers of Mathematics. Even though this aspect of governance must be considered "top-down," reliance on the various professional organizations to determine what ought to be learned will result in a "bottom-up" effort where participation at the local level across the nation is guaranteed. No certification of these standards will be required, since it would be redundant -- and self-serving -- to

have the governing body essentially approve whether its own oversight of the development process has been acceptable.

Step 2: Intermediate Building Blocks

Since there are inadequacies in the Council report as it relates to the purpose for new assessments, the governing body would be required to convene and decide upon the primary purpose for which new assessments are needed. AAP requests that the Subcommittee bear in mind that if there are multiple purposes ultimately found to be warranted, more than a single new assessment would be required because one test cannot usually perform more than one purpose. Moreover, since there are strong indications that performance-based testing does not meet high technical standards, additional research should be undertaken under the direction of the governance body to determine how new forms of assessment will meet technical standards.

Step 3: Intermediate Resolution of School Delivery Standards

The new governing body should also be responsible for collecting data on the implementation of the delivery system. Regardless of whether the system has a federal component, content standards alone are not sufficient to produce quality instruction and educational improvement. For some critical period after "world-class" standards are developed, it will be necessary, at a minimum, to focus on obtaining and information on: (1) how many states and/or school districts have adopted those standards; (2) what types of instructional materials have been produced that incorporate the new standards and what is the level of use for these materials in the states/districts; (3) what percentage of teachers in each particular subject area have been trained in the new standards (by state, district, and school); how many students are being taught using the new standards (by state, district and school); and (5) other data on resource availability or distribution, such as the number of computers in each school, the amount of teacher training in assessment methods, etc. If we expect to hold schools and school systems accountable, then these tasks are imperative.

Step 4: Intermediate Assessment Activity

It would be a mistake for this Subcommittee to believe that nothing important is going on today in assessment. As I mentioned earlier, publishers, as well as assessment centers and various consortia, are working on different projects that could hold a great deal of value in achieving the ultimate goal of improving student learning -- once "world-class" standards are in place.

The Council, should it be created, or the Goals Panel, should also be looking objectively at the various innovative programs now

underway or being planned that "push the limits" of assessments, both as to the purposes and the technology. Congress should be advised as to how successful these programs really are in improving student learning, as well as evaluating them on technical factors and cost-effectiveness. Will they do what their advocates claim? How successful will we be in incorporating national standards in our tests, adding performance assessments, and including higher order and other thinking skills in our multiple-choice format tests? Alternatively, this objective research could be conducted under the auspices of OERI.

While the standards are being developed and consensus reached, while textbooks and other instructional materials are being prepared, and while teachers are being trained, let it be an opportunity to provide Congress, educators and state and local officials -- as well as the public -- with objective and reliable information. Let us look before we leap!

For our part, publishers stand ready to contributeto these efforts in building a body of knowledge to enable Congress to act in a fully-informed fashion when called upon to do so.

Step 5: Reports to Congress

Congress should establish appropriate timeframes for the governance body to report, both on an interim and final basis. Those dates should be set at the earliest possible periods, given the importance and complexity of the tasks to be performed. Such reports, and their thorough evaluation, must be the precondition for any Congressional action.

Step 6: Final Confluence and Action

The last step in AAP's recommended process can occur once all of the data and information has been collected on standards and their implementation, on school delivery standards and where resources have been committed to enable students to achieve the new standards, and on the status of assessment research. When Congress has received all of this information, then it will be ready to act, if necessary, on recommendations about new assessments and to determine the appropriate role such assessments should play in the education system. Congress should direct the construction of properly focused and developed assessments only when it has all this information.

We submit that this approach is the best way to avoid the problems and shortcomings AAP and many others have identified with the NCEST report. Unfortunately, there is no way to short cut this process without jeopardizing the very school children who are the subject of this effort. Assessing students must await the successful and complete implementation of "world-class" standards, but commercial publishers are committed to building appropriate

tests for the intended purpose(s) once those standards are in place.

CONCLUSION

Mr. Chairman, I am sure that if any of us could wave a wand and make students learn more, become better functioning adults, become more capable of joining the global work force, and be better citizens, we would do so. The problems with our educational system are not that simple and the answers not that easy. Certainly, it is not a situation where the tool of assessments of how students are doing can somehow be transformed into its own solution -- and that is what the NCEST report erroneously tried to suggest. But Congress can require that a carefully detailed and responsible path be constructed that should lead to those desired outcomes.

Let me close by reiterating my four major points:

1. Assessments -- even those built upon "world-class" standards -- will not alone achieve the goal of improved learning.

2. What we want students to know and the opportunity and resources to them it must be provided before a national assessment system is put in place.

3. All assessments used beyond the classroom must meet the same high standards of validity, reliability and fairness.

4. There is no need for a national, centralized "top-down" governance system for assessments. It would be duplicative, costly and could stifle innovation and competition.

I appreciate the opportunity to present the views of the Association of American Publishers on this important subject. I will be happy to answer any questions you may have at this time.

Chairman KILDEE. Thank you very much, Dr. Kean.

I'll address my first question to the Governor. Governor, the Council recommends voluntary national student performance standards, but school delivery standards are developed by the States. I know that was greatly debated in the Council.

Would you personally object if this committee required the new national entity, the National Education Standards and Assessment Council, to establish voluntary national school delivery standards?

Governor ROMER. Congressman, I know you saw me chair that group. There were times in which there were various points of view that we had to put together to get a composite. I'm the kind of a guy that when I make a deal, I'm always with the deal. Do you know what I mean?

So, I come to the table really wanting to defend the report. Now, you also know that it was not my initial position that they be developed—it was my initial position that they could be worked on, you know, simultaneous with the others at that level.

But I don't want to play games with my colleagues on that group. I honestly came to the table knowing it's very difficult to move this forward in America. It's very difficult. You need to sometimes make some compromises in the way you go.

So, I just want you to know, and let me state it exactly, that we need to have delivery standards developed. I think that the Council ought to have the responsibility to see that they're developed. But I would hope that if it does vary from the report, that the development of delivery standards would be fully cognizant that there is more than one way to deliver, and that we ought not get a master recipe at the top which we imprint on everybody.

Do I communicate effectively?

Chairman KILDEE. Yes, and I understand your position. Very often Adlai Stevenson, when he was ambassador to the U.N., had to be ambassador and represent the position of this government and very often not his own views.

I didn't mean to be unfair in asking that question, but I do respect you not only as chairman, co-chair of that Council, but as someone who has really worked hard in education and as the Governor of Colorado where you worked very hard also.

The reason I asked the question is that I know by looking at the people, particularly on this side of the aisle here, and not just exclusively though, is that they raise some very serious questions about having national student performance standards without having some, at least symmetrical, same level national delivery standards because of some of the reasons I gave prior.

You can predict in certain situations just how the students will perform, depending upon, as was pointed out here, the standards/resources that are available in those schools.

Would anyone else want to comment on that question? Yes, Dr. Kean?

Dr. KEAN. Well, I would agree with your concern, Mr. Chairman, and caution that without providing for school delivery standards and then assessing, one runs the potential problem of what researchers refer to as the measurement of nonevents. If that is the case, the question is for what? What are we assessing?

Chairman KILDEE. Particularly, if we're going to have risk, high risk, medium risk, any risk, should the deficiency of the school follow that student all the way through his or her life if the student isn't able to perform certain functions or is not that good at math, certain fields that he or she may not feel they should enter?

But if the school has failed—I guess we have to ask is it the student who has failed or the school who has failed. In many instances, I think it's the school.

Governor?

Governor ROMER. I had a very interesting experience sitting here this morning. Let me just bring a couple of elements to the table, Congressman. I think you know when we formed the goals panel, we had to carry that kind of tenderly across a mine field to get it going. Then we recognized that it wasn't any good to measure something if you didn't know what it was that you were measuring against.

So, we knew we had to turn to standards. I think there's some agreement here at the table on standards. I want to simplify this. I'm just kind of a guy from the country coming to this great city and say folks, we want to do the best job we can in math. Have you got anything to share with us on math and on science and on English and history about what is really world class?

Now, hey, we don't want you to force it on us, no. It's not mandatory. But could we just come to town and say give us the best you've got. Don't even develop it by "the best and the brightest;" do it in an iterative process throughout America.

That ought not scare people. I mean, as I listened to Mr. Dean and Mr. Koretz, I thought there was some horrible apparition that was about to devour values in education. Almost the caricature of this report, which I heard here, bothered me.

But I just want to bring it down to simple language. I think, Mr. Kildee, that if we do not create some mechanism that continues this work, I know where it's going to fall. It's going to fall back in the Department of Education. I think that is not the only place that this work ought to occur because we do not want to arrive at a ministry of education.

I know the politics of America and education well enough that if we can have some ongoing participatory body and standard, it is a good thing. So, I'd urge Congress to continue this work because I think the 6 months we spent together, we helped forge some of the debate in this country in a healthy way as to where we ought to go.

The next thing is on performance standards. The high stakes is driving a lot of paranoia and fear. I grabbed the report and began to read it, and I said did we say all that about high stakes? I don't think we did. I think that we ought to be not fearful of proceeding appropriately to improve the way in which we assess.

I never intended nor did I think that report intend that we use the product of our work in a dumb or a stupid way by using it for high stakes in the wrong way. I'm just curious about that.

Chairman KILDEE. You want to comment?

Mr. KORETZ. Well, yes. I'd like to avoid getting into a debate about what exactly the report says, though I do have my briefcase page references for the high stakes. The report lists things like screening kids for jobs. That's very high stakes. You could lose a

job if you don't do well on a test. That's about as high stakes as you can get.

It lists quite a number of high stakes, but moreover, research in the last 10 years has shown us that you don't even need to go anywhere nearly that far before you start to get perverse affects in schools. In one large district where I've recently finished a study, we deliberately picked a test that was not the one that had formal sanctions attached to it; it was simply one that was used for job boning and publicity.

Even on that test, the inflation of test scores was more than half an academic year by the end of third grade. The reason is, teachers had gotten the message loud and clear that they would be rated on the basis of how kids scored on that test. That's all it takes.

Now, when the Governor says we don't want to do things in a dumb way, I can't imagine that any of us would disagree about that. The question is what is the sensible way to use tests with high stakes. The Council report doesn't provide one.

Let me go one step further. The problem is not only, in my mind, that high stakes testing creates all sorts of deleterious side effects; the problem is that it simply has not worked as the primary engine of raising standards and improving performance in the many times it has been tried. I don't know why we'd want to try it again when it hasn't worked before.

Ms. ROSENBERG. Let me just make a comment. While there may be some people who want tests, a new series of tests, to drive reforms, there are other people who subscribe to the NCEST report and there are many different ways of reading it. Like I said, this should be the beginning of the conversation not the end.

But there are other people that have a great interest in the standards part of the argument in reaching some sort of consensus, not in excruciating detail but some sort of a consensus on what we want our students to know and be able to do and the kinds of performances that we value.

We are obsessed on the test issue. There are some good reasons to be so obsessed. We assume because everything that has gone wrong in the past we will just repeat those mistakes in the future, and that we are powerless to do anything about them.

I simply reject that point of view. The issue is not so much the tests, except insofar as there are good tests and there are worse tests, the issue is the accountability systems that we design. There are a lot of people up here now who have spent a great deal of time criticizing the present tests, criticizing the fact that curriculum is now a line to the tests rather than the other way around, they are, in effect, defending the status quo.

Mark my words, the status quo is not going to wither away of its own accord. All these local initiatives and State initiatives that we hear so much about, I'd like to know where they all are. We can all name the same ones. What dynamic is there in the system to make that spread? There is none.

That's the way we should be reading this report, to change the dynamic of the system so we stop wringing our hands, so we stop just having the handful examples to point to. When their superintendents leave, those examples will leave with them.

Chairman KILDEE. Mr. Pastor?

Mr. MARTINEZ. Will the gentleman yield?

Mr. PASTOR. Yes.

Mr. MARTINEZ. On that last statement you made, when I read—and I'm not "antistandards" and I'm not "antitesting" kids to evaluate where they are, where they should be. But Governor Romer just mentioned that he wondered if that's what they said in the Council. He was reading the report there and Mr. Koretz was referring to his packet in his briefcase. But I've got the page right here in front of me. I marked it because it scares the hell out of me.

It says the Council finds that the assessments eventually could be used for high stake purposes. It says they could eventually be used for high stake purposes such as high school graduation, college admission, continued education and certification for employment.

That's why we're into the debate about this test and why the OTA report talks about it in every sentence and every other word what the testing could do and what it shouldn't do, because that is uppermost on a lot of people's minds.

Yet, we know, for example, that there have been people misplaced by tests that were given. The Council of Chief State School Officers found that a large number of normal children who don't speak English as a first language were being misplaced and miseducated in special education, in which they really didn't belong.

I mentioned earlier my experience of being put in a speech correction class with a bunch of young people that stuttered. Due to those and other problems, they've sure got causes to be paranoid about what the use of the tests are. My concern is what is going to make whoever is setting these things accountable to the public and accountable——

Mr. KORETZ. May I respond to that? Congressman, that is precisely—and I think I speak for the other three people who are with me, but let me just play it safe and answer for myself. That is precisely what bothers me. The accountability is for kids and teachers, but apparently no one else.

I think that when we propose dangerous medicine, we ought to be required to test it. I recently went through an exercise with two colleagues looking at innovative performance assessment programs around the country trying to find even the most rudimentary evidence about the quality of the assessments or their effects. The answer is it's not there.

We're doing the same thing we did in the 1980's when policymakers said boy, let's test the heck out of kids using standardized tests. Everybody started making teachers and kids accountable for it. They said there's no time to waste, so let's not find out whether it works. Well, it didn't.

I think we have an obligation. I don't think this is a matter of obsession or paranoia or other character flaws of the critics of the NCEST report. As a psychologist, I don't think that's really the issue. I think the issue is public accountability.

When you do something that's potentially risky and relatively untried with nine million kids, we ought to find out what we're doing.

Governor ROMER. Could I reply to this? I find it very interesting the interpretations that have been given to this. First, Congressman, high stakes or some stakes are already given to the methods that we use to test now. Now, I'm into this game not because I want more testing, not because I want to increase the stakes that are put on testing, I want to get better tests, pure and simple.

Now, the argument here—I wrote a note when Mr. Koretz was testifying. The conclusion I got was tests are bad. Tests are bad. It was the simple way of summarizing in an unfair way what he was saying. Look, we already have an awful lot of testing out there. I am highly guarded against the wrong use of tests.

If I had had more time this morning, I would have put up on a chart the users of the test and who should be using it. It would have begun with the student, and the teacher, and the parent, and then last on that list would be national or State.

But let me bring this back to the real world. Congressman Kildee, I hope we pass this bill with this in it. I hope we can begin to create something that continues this momentum towards better standards, period. Now, if everybody wants to shy away from tests at all, let me tell you what we'll do. We're already doing it.

Colorado and 16 other States are associated in a private venture called the New Standards Project. It's the best I know out there in the country, which is working on new forms of assessment. If we don't have anything to work in a national way, we'll create something on an association of States to work at it.

Now, let me tell you what we're trying to do. We're trying to find a way, such as in English, that we really get teachers much more involved in the assessment in which we use very much a more enriched and comprehensive way to test such as essays. We understand that there's a time and a cost involved here. We also know what it is you're trying to reach in good testing for English.

I, frankly, am appalled. I was appalled with the testimony of Mr. Kean representing the textbook publishers. I kept trying to figure out what is his motivation for what he is saying. Was it to protect the economic interest of the tests that are now on the marketplace?

I think that all those textbook publishers have got a great market here if they joined with us and find a way to produce better tests. I just don't understand, quite frankly, why the high powered rifles are aimed at what I think is a fairly modest constructive step to get something better rather than something worse.

I find great ghosts and caricatures are being drawn of the consequences of this which I just don't think are justified. This is a voluntary exercise. It is a voluntary exercise. If I in Colorado don't like it, I don't even have to submit anything to it.

Mr. PASTOR. Mr. Chairman?

Chairman KILDEE. Mr. Pastor?

Mr. PASTOR. If you could have seen your face, Governor, I could see that you were appalled many times during this testimony.

Governor ROMER. Well, I get appalled too many times in too many places in this city, as you well know. I'm trying to learn to do better.

Mr. PASTOR. Before we get to the assessment, I think we need to get a consensus of what the standards are going to be. I think get-

ting that consensus is going to be very time consuming. So, let me get to that question.

In developing the standards, I've heard words like world class. So I'm assuming now that in the past when we created standards, it was basically for internal use. We want a person educated so they can find employment in a factory, farm, et cetera.

So now we're talking about competing or a standard that would make a person be able to compete worldwide. So, I'm assuming that's the term world class. I heard Mr. Kean. He said that the process ought to include people who have experience in curricula, that those are the people who can develop these standards.

My question to you: Would we develop these standards—and I hear top down, bottom up. That will be an issue. But do we compare ourselves now to other industrialized nations to develop these standards, or do we develop them with what we think ought to be the standards? What would be the process?

Governor ROMER. Let me tell you what I think ought to be the process. We could take math or science or whatever. Let's take science. We ought to develop an initial draft of what we think this is using the best minds of America, using the people who are in the subject matter area.

We ought to take that to everybody in America we can get to listen to it. We ought to take it to every science teacher, parents and to the business people of this country and say what do you think about this initial draft? Does this really prepare people for the kind of jobs you see that we need out there?

Then that response ought to come back to a central location. It ought to have a redraft and it ought to go out a second time. That time we ought to have a Monday night set aside and we have 2 hours interactive TV with every cable outlet in America where we can hopefully get all the parents focused upon okay, this is really important.

My son or daughter may or may not have a job, whether or not they're taught the right thing. I need to know what it is they think these standards are. Then, it ought to come back a second time at the national level and get another draft. Then, that ought to be made available on a voluntary basis.

Now, let me tell you why it's important to have that process of participation. If you really want to move this country, you're not going to move it by writing standards and putting them on some shelf. You're only going to move it when parents know this is what my youngster needs to know and be able to do to be successful.

We have got to do this in a very participatory way. You see, this is the problem. We don't have a mechanism right now to do that other than the Department of Education. That is not the right mechanism because it becomes too federated, too Federal.

So, what I'm appealing for here, let's create something that together, and I mean together, I mean States, Congress, and the administration, can engage this whole Nation into an understanding of what really it is we ought to know and be able to do.

My frank suggestion is we take a subject matter a year. The next 12 months let's take math. Let's take it to everybody in America. Let me just say it would be great to have a Monday night in which we had every parent in America with a little notebook in front of

them and saying this is 8 pages or 12 pages of what my youngster ought to know and do in math. I need to react to whether that's the right standard or not.

Now, most American parents are going to be blown away by that. But you understand the education that would occur for parents if we got into that? That's my concept.

Mr. PASTOR. Who would have the initial responsibility for a draft?

Governor ROMER. Okay. In math, it was the National Council of Teachers of Math. In science, we've already put together a group, and they have $1.5 million out there. That was started by the Academy of Science and I think another group.

You see, these things are already happening, Congressman. I am desirous that we have some oversight as to how they happen because I don't have input, quite frankly, in the way they're happening now. I, as a Governor, would like to have input.

I think there are better ways and worse ways to do this. But the key to it is it isn't enough just to get experts to write the report; we've got to involve all of America in understanding what it is.

Mr. PASTOR. Anybody else want to comment on that?

Mr. KORETZ. I would like to go back to the last iteration. We've progressed from paranoia and an obsession to questionable motives and lack of reality testing. I think we ought to address these arguments on their merits and I think we haven't done that.

I wanted to clarify one point, and one point only. None of us, of the four of us who prepared the testimony that I presented, argued that tests are bad. Speaking for myself, I have been involved since the very first day in Vermont's efforts to design portfolio assessments, which is part of the new standards project you refer to.

What we oppose is using test-based accountability as the primary engine or reform with insufficient knowledge and insufficient precautions. We oppose that because the evidence overwhelmingly suggests we ought to. I am disturbed by the fact that as this argument has gone on, and it's gone on long before today, I have found it very hard to find people on the other side of the argument who are willing to say no, your reading of the evidence is wrong. Let me show you the evidence that you've missed.

In fact, in debates in the assessment task force of the National Council, of which I was a member, as was George Maddaus who helped write this with me, I explicitly asked people on the other side to come up with evidence to support their position, and they weren't able to do so.

I think the simple fact is we have a lot of experience that says we're barking up the wrong tree. That doesn't say we shouldn't have better tests. That doesn't say we shouldn't have higher standards. It doesn't even say we shouldn't have, if you want to go that far, a national curriculum.

It says we shouldn't use test-based accountability with half designed and unevaluated tests as the primary mechanism of reform. I think that message is fairly clear.

Mr. PASTOR. Mr. Chairman, just to continue. Mr. Koretz, don't you think that the assessment, the testing, would only come after the definition of what is an educated person through some kind of standards?

Mr. KORETZ. Well, Congressman, I would hope so. I'm a little fearful that it might not. Originally, when the new standards project, which Governor Romer referred, got going, that sequence was explicit in their plans. We will develop standards. We will develop syllabi. Then we're going to worry about tests.

But once the political pressure started to build for doing it fast—somehow it's going to be catastrophic if you don't do it faster than we've ever done it before—the sequence seems to have gotten muddled. At least in my mind, I think the only safe way to do it is to first determine what it is we want kids to know.

Mr. PASTOR. Well, I agree with you because you only test because you want to make sure that the people are learning what you're asking them to learn.

Mr. KORETZ. Exactly.

Mr. PASTOR. That would be my position until we have some standards. That's why I was interested in what process do we implement to assure that we get national standards that will meet the needs of America? Then, once we have those standards, then how do you evaluate them?

I was interested in the process because the process should also be timely. But if you involve parents and try to make it inclusive to get a consensus, sometimes that's very difficult. But obviously, there has to be public input or it wouldn't be accepted.

Mr. KORETZ. I agree. I think Dr. Kean was very sensible when he said it should be as rapid as is feasible. In Vermont, where we started with the NCTM math standards as a given, we are now on our first full trial year of an assessment and only our first year of evaluating that assessment. Vermont is one of the few States that actually is committed to an evaluation.

I think, frankly, if we want to do it right, we're talking about something that's going to take quite a while. In Vermont, we've been blessed by the fact that the political climate allowed the program to stay pretty much in tact through three Governors. I don't know if that would be true in other States, but it's taking a long time to carry out.

Mr. MARTINEZ. Would the gentleman yield on that point?

Mr. PASTOR. Yes.

Governor ROMER. Could I add one more thing here?

Chairman KILDEE. The gentleman wanted to add something to that point.

Governor ROMER. Maybe we're getting closer to the core of the matter. You know, I'm not a professional educator and I'm not a psychometrician. But I've spent a lot of time on this subject. Mr. Koretz says the problem is we're aiming at test-based accountability. I'm not after accountability in this thing. I'm after improved instruction in the classroom. That's really what I'm after.

Let me tell you what we're doing in Colorado. I'm aiming at the point that standards and tests are not separate; they are related. They are related. Let me tell you what we're doing in Colorado right now. I have a team of teachers arising out of the new standards project that are working on items. These are items in English and math.

What is an item? It's an illustration of whether or not you know how to do a particular function or an educational process. Why are

we developing items? Because we know that we are trying to figure out what is the standard we are reaching for.

Also, we're trying to figure out a way how we test whether or not a youngster is reaching that standard. So, the process of developing a standard and developing an authentic way of assessing whether you've reached it is intertwined.

They are not separate. I just have to differ with my colleagues at the table that you "go develop standards" and then you draw a double line and then you develop assessments. Standards and assessments are intertwined. You do not need to have a lineal process of 7 years to do this. You can do these together.

Now, I'm not taking just from Romer here; I'm talking about some of the better people I find in the country who are working on this through the new standards project. So I guess I just want to conclude and say that, Mr. Koretz and others, Congressman, I'm not after accountability; I'm after improved instruction in the classroom.

A good test can be very creative in doing that. It doesn't have to be there to put a hammer on somebody.

Chairman KILDEE. Ms. Rosenberg?

Ms. ROSENBERG. Just a few points. One is we fully agree that the standards need to precede the assessments. I would also say we should start a process of getting rid of some of the current idiotic tests that we have that depress teaching and learning.

The other thing is that high stakes has different meanings to different people. There has been a gross sort of caricature of the meaning of high stakes. We are all the products of systems in which tests counted, not for everything.

When we try to get into college, test results were looked at, grades were looked at, teacher recommendations were looked at. There were a number of things that were looked at. Tests were one of those things. No one is talking about determining an entire life based on taking one test.

While there are some people, it is true, who are talking about the results of a test for job screening and go/no-go decisions on college, that is not the only meaning of high stakes. So, again, I urge you to keep an open mind about the meaning of this.

Second of all, if national standards and examination systems led to all these horrible things and suffering children and pernicious consequences, then I do not understand why all of our competitor nations have had such a system for years and their children are not miserable. They are not walking zombies. More to the point, they achieve at much higher levels than our students do.

I think that we have to get a little bit realistic about what the meaning of national standards and examinations are rather than this romance about this creativity at the local level and all the terrible things that could happen to youngsters on the virtue of tests.

There's no question that how we design an accountability system is critical, absolutely critical in shaping the behavior of the adults in that system. But we are running a public system, and the public has the right to know.

There has to be room for subjectivity in that system for judgment. That's why we say that tests at one point in time shouldn't be the sum total. A teacher who might have looked at a test result

and then looked at you wouldn't have misplaced you, such as the tyranny of tests right now. But unless we have a testing system, we have absolutely no way of knowing how our students are doing.

On the equity issue, let me point out to you that it is not until that we had tests that we knew just how miserably we were doing with minority students and had some concrete evidence to start moving some people along. Now, have tests been misused? Absolutely.

But once again, we've learned a lot. I would rather put our energies in designing accountability systems that don't suffer from those defects than just throwing our hands up and reifying the status quo.

Mr. MARTINEZ. Did she use up all of my time?

[Laughter.]

Chairman KILDEE. Let me ask you a summary question. It's all been, I think, mentioned, but a summary question and your comments on it. What is the priority of time and the priority of importance between standards and assessment? That's kind of a philosophical question if you mull it over.

I, myself, worry about lack of school-based standards. Governor Romer knows that. But what is the priority, or do they, of necessity by nature, run parallel roads at the same speed?

Governor, do you want to start?

Governor ROMER. Can I take a crack at that? First, on school-based standards as against the student standards, I think they need to run simultaneously. We need to be very thoughtful about school-based standards because we need to have a whole lot of innovation and try some new things.

In Denver, we're trying some new things. We want to have school delivery standards written in a way that we leave room for more than one way to do it. But I think that the effort ought to go simultaneously.

Now, in reference to assessments and standards, even though I have very reputable advisors who say they should also be done in parallel, I, in my own gut, feel like there is a standard thing you do first and the assessment comes along a little bit after. But they are not totally linear They kind of overlap.

Just in specific terms, timing, I would hope that in the next 12 or 18 months that we can really get this—well, I think we can get the process beginning immediately, but hopefully we can complete the first round because this is a dynamic changing thing of standards in the next 12 to 18 months.

Then I think assessments, because of the complication of that, is something that's down the road after that, even though we work at them parallel.

Chairman KILDEE. Yes, Dr. Kean?

Mr. KEAN. Just a word to add to that. Until one decides what the purpose of the assessment or the assessments are, then you can't build them. You need to know what information you want. There's not a right or a wrong, and there's probably a variety of different types of information that will be useful.

But that question must be asked and must be addressed in order to build proper assessments. If it's information for classroom teachers to use in improving the quality of instruction, terrific. But a

certain type of assessment, a performance-based assessment in certain instances, is appropriate for that.

But when we commingle and when we list a half a dozen or more different purposes for an assessment, then we're dooming ourselves to failure.

Chairman KILDEE. Yes?

Mr. KORETZ. I would like to pick up on that. I certainly agree that standards have to come first, that it just doesn't make sense to try to design a test until you know what kids are supposed to know. But if the question is how quickly can we design tests that will monitor whether or not kids are learning what's in the new standards, I think the answer is that can be done reasonably quickly.

We have a national assessment of educational progress as a model. It can be modified to incorporate new kinds of skills, new kinds of content, and so on. That's not the sticky wicket. The sticky wicket is trying to use tests to do all these other things that the Council's report suggests and that evidence suggests don't work very well.

So, if I were to pick a sequence, I would say start on standards immediately. Shortly thereafter, start on refining tests and monitor how well progress is being made toward those. But much further down the road, with a great deal more research and serious evaluation, experiment with tests for other purposes such as reforming instruction in classrooms or holding people accountable.

Along those lines, I just want to point out Ms. Rosenberg really missed the point in terms of the information tests provide us. We know that kids aren't doing as well as we want them to because of tests that have not been used for accountability and that have not been taught to, specifically, the national assessment.

If you look at tests that have been used for accountability, many of them say things are wonderful. That's part of the problem.

Mr. MARTINEZ. Mr. Chairman?

Chairman KILDEE. Yes.

Mr. MARTINEZ. Could I get some time?

Chairman KILDEE. Yes. You've got plenty of time. I've got all day. I really don't, but——

Governor Romer, you raised a good point. We don't want to inhibit with school-based standards innovative, creative things that they place. I wouldn't want that. For example, I believe something I like to encourage is site-based management.

We wouldn't want to discourage that by school performance standards, but if we take the explanation that Ms. Carson gave us, standards/resources, now in any school, if our standards were based more about what the resources of that school should be, I think then we're not getting into curriculum.

We're not getting into inhibiting site-based management or creativity, innovation. But I think maybe we could look on some standards as what should the resources of that school be, and what really—I know why some of the States are a little concerned about this.

They hate to be told what resources should be used in their schools. We saw that when the Council met. But since these are voluntary anyway, to say this is what we feel the resources should

be in a certain State, a certain school system, that would not destroy creativity, innovation, site-based management.

But maybe standards/resources, as Ms. Carson testified, would be the way to address the school performance standards of making sure they had at least a floor of resources, a minimum, at least, below which you cannot really perform well. That might be the standards that we could talk about.

I do know some school districts that just do not have enough resources to really be fair to those students who are going through that system.

Governor ROMER. Could I reply? I think that the resources—we need to also, though, define it. In Colorado, for example, Vale is a very expensive community. The resources to hire teachers and to find accommodations are very much different than my old home town of Holly, Colorado. So, there's a standard even within Colorado that we need to know.

We set our school finance formula based upon cost of living, in fact, in certain parts of the State. Now, it's a fairly rational one, but I've thought a lot about what are the indicators of good practice. Is it site-based management? Is it the number of students in the classroom? Is it the length of the school year? What are the "delivery" standards?

I think we need to work at them. But the difficulty is that there's going to be a lot of variety, a lot of variety in that area. Let me go back to the pole vault. You can use the flaws were a flop or you can use a whole lot of other ways to get over the bar.

But at least I think it's accurate that we ought to report to each other this is what the bar level is as they use it out there in the world.

Chairman KILDEE. We would not want to bring down those who are performing well to the level of those who aren't doing so well. I think the idea would be to pull up those who aren't performing well. Maybe one of the reasons they aren't performing well is they don't have the right resources. So, I would hope that we would try to pull people up. I can recall when I was teaching. The janitor in the school I taught was making more money than the teachers. We would have teacher meetings and they were complaining that Mr. Smith was making more money than we were.

So, I finally said we're not going to help ourselves by pulling Bill Smith back down to our level. Our job is to get ourselves up to that level. So, I think we should at least maybe talk about some minimal standards of resources in schools.

Those standards where you have special problems, you might need special resources then too; right?

Governor ROMER. Yes.

Chairman KILDEE. When you have a lot of poverty, a lot of crime, homelessness, drugs, you might need some special resources there, too. So, very often you can say well, a certain school has advantages and resources. Some have disadvantages and no resources.

But maybe we can put some extra dollars—we do that with Chapter One; right? We drive some extra money into those areas where there's some special problems. That's the whole theory behind Chapter One.

But I do think that the idea that I like is that at least on standards, we should at least look at the idea of encouraging the use of adequate resources in all school districts.

Governor ROMER. Yes.

Chairman KILDEE. Yes, Ms. Rosenberg?

Ms. ROSENBERG. Let me endorse that point of view but add another point to it. Frankly, given how tricky the school finance issue is in the States and local districts, not to mention at the Federal level, I think that the process of establishing content and performance standards, what we want our students to know and be able to do, is an important leverage point for moving that resource discussion.

Right now it tends to take place either at a level of abstraction or merely the relative power between the haves and the have nots. So, you have situations like the New York City schools getting far less from the State than suburban schools, even though the needs there are far greater.

It's just a discussion in the abstract about the cost of bureaucracy, whether the money is being spent well or not. If that kind of discussion can be framed around the issue of standards and where different populations are relative to that and opportunities to learn, I think we have a far more legitimate way of moving the public on the issue of equality of opportunity.

Chairman KILDEE. Thank you very much.

Mr. Martinez?

Mr. MARTINEZ. It's late and I know you've been here a long time, so I won't keep you very long. Incidentally, I just wanted to tell you, Governor Romer, that I'm originally from Colorado. I was born in Walsenburg, Colorado.

Governor ROMER. All right.

Mr. MARTINEZ. My daddy went, with all of seven children (and there were three more born in California), to the Golden State—where the streets were "paved with gold." We found out they weren't. My dad used to say a lot of things, and sometimes listening here to the conversations the things he said ring back to me.

I know you, describing yourself as an old country boy, would understand "the horse before the cart," "Figures don't lie but liars often figure." In that, maybe I do have a paranoia about testing. I'm not against testing, but I would rather see testing result in evaluation and assessment that results in the best end product: the ability of the child to learn.

In any business that you have, the end product and quality of that product is the test to your ability to do the job well—and thereby the accountability that you talk about. But we've had, in the time we've all been alive, the Vietnam War. Do you remember the body counts that were indicating that we were "winning the war"?

In fact, we weren't winning the war despite those empirically-based body counts. They weren't accurate. They didn't give us the right information.

Right here in Prince Georges County schools, we had a gentleman that gained national reputation for the "improvement" of the county's schools.

Remember what happened? It seemed like what he was doing with the scores to show his achievement wasn't accurate because—he got a better job. He became famous, got a better job. Subsequently, they administered another type of test and the scores plunged.

The thing that caught my attention is that in those first results that showed great progress, minorities were shown to have the greatest progress. Conversely, when the tests showed the real results, the minorities plunged the worst.

I'll go back to Jamie Escalante who gained national fame. Stand and be counted; remember? Jamie Escalante is no longer the teacher teaching that class. That class has never done as well. So, you leave me with a lot of questions.

We've got to get away from that "Vietnam body count" syndrome. When we talk about single measures, we fail. You talk about volunteerism. Well, that's fine except that you also talk about local control and developing the delivery systems.

But when you set up a national body and a national standard and a national testing, and you have people that volunteer into this program, who is going to control? The national level because they set the standards or the local level?

I mean, those things leave me in a quandary. It's not that I don't want to see progress, not that I don't want to see what you want to see: the end result, the best benefit to that child.

The other questions that are not answered here are the questions we and Mr. Kildee talked about: the inequity of school financing. In California, we had the Serrano vs. Priest decision that was supposed to provide equitable funding for all schools and all children.

It still hasn't achieved that on a State level. How do we achieve it on a national level? Unless we address those problems of opportunity to learn in accordance with all this and while we're doing all this other, we're not going to achieve what you and I both want.

It leaves me with a lot of unanswered questions. I have a tendency to understand that reports can be interpreted, as you said, Ms. Rosenberg, the way the person is reading them. We all read the same thing and interpret it 10 different ways.

What we have to have is a universal understanding of what we're talking about. If those words are not accurate, then we ought to make sure that the report is amended to be accurate as to what you're saying about "high stakes" purposes.

How do you, from a local government, respond to the questions that I've just raised about who is going to control?

Governor ROMER. Congressman, let me say, I have opposed a national test. I have opposed exit exams for high school. Whenever I talk about testing, I talk about the primary purpose is to improve instruction of the classroom. I talk about beginning at the bottom, the fourth grade, because I want us not to use it in a high stakes way.

I have been consistent over a period of time on this. But let me tell you where we are. This Nation made a great deal of commotion about arriving at educational goals, you know, the six national goals. The president and 50 Governors said that.

They created this panel, and I was the first chairman of the panel. All the panel wanted to do was to measure, measure, meas-

ure, measure. You and I are from rural Colorado. You can continue to weigh a steer, but if you don't ever feed the steer, the steer doesn't get any fatter.

Now, I wanted to move off of this measurement annually into doing something about reforming education. There is about 20 items of reform. I don't believe that testing and standards are the key to all of that. But I think they are an ingredient of that.

I think what we're struggling with here today, and to use your fear, and that is if we create a national group, we don't want them to go in the wrong direction. I think you can put those safeguards in. I think there's unanimity here and the table of standards is a good thing.

Let's get a mechanism by which we could go out and develop them together. Then, let's talk about assessment. Put in there that assessment is something that's going to be State by State. It's local. Maybe we do R&D. I would love to have some R&D with assessment.

Third, if you find that at this point in time a certification on a voluntary basis is a scary thing, knock it off. Knock it off. Let's get ourselves going down this path and we can readdress it.

What I fear is that if we don't start working on the substance of reform and we just keep measuring like we are in the goals panel, it's kind of like weighing a steer every other day to see if it's gaining any weight and you're not feeding the steer.

Mr. MARTINEZ. I agree. Thank you, Governor Romer. Thank you, Mr. Chairman.

Chairman KILDEE. Anyone want to make any closing statement or any comment? Anything further?

[No response.]

Chairman KILDEE. I want to thank you all. I think we've found some consensus, more consensus on standards, some diversity on assessment, all of which has been very, very helpful.

I enjoyed serving on the Council. That was a good education for me also. I really got educated here again today, one of the good advantages of serving in Congress.

All of you bring to the table expertise and deep abiding love and interest for education. So, we're the beneficiary of that. We thank you for that. We'll keep the record open for 2 additional weeks for any additional testimony.

Thank you very much.

[Whereupon, at 12:15 p.m., the subcommittee was adjourned.]

OVERSIGHT HEARING ON THE REPORT OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING

WEDNESDAY, MARCH 18, 1992

HOUSE OF REPRESENTATIVES,
SUBCOMMITTEE ON ELEMENTARY,
SECONDARY, AND VOCATIONAL EDUCATION,
COMMITTEE ON EDUCATION AND LABOR,
Washington, DC.

The subcommittee met, pursuant to call, at 10:05 a.m., Room 2175, Rayburn House Office Building, Hon. Dale E. Kildee, Chairman, presiding.

Members present: Representatives Kildee, Sawyer, Lowey, Reed, Roemer, Goodling, Petri, Gunderson, and Pastor.

Staff present: Susan Wilhelm, staff director; Damian Thorman, legislative associate; June Harris, legislative specialist; F. Jefferson McFarland, legislative counsel; Steve Boyarsky; and Lynn Selmser, minority professional staff member.

Chairman KILDEE. Good morning. This hearing of the Subcommittee on Elementary, Secondary, and Vocational Education, an oversight hearing on the report of the National Council on Education Standards and Testing will come to order.

The National Council on Educational Standards and Testing has recommended that the Congress enact legislation to provide for the development of a national educational standards and a national system of assessments.

In addition to Secretary Kearns' testimony on the National Council's recommendations, the subcommittee will hear from three important perspectives. It will first examine the current and potential effects of standards and assessments on children of limited English proficiency and on exceptional children.

The committee will also hear from Dr. H. D. Hoover, author of the Iowa Basic Skills Test, on his response to the council's report. Each of these perspectives will greatly assist the subcommittee in determining what the proper course of action should be. At this time I would like to recognize Mr. Goodling or Mr. Gunderson, either one, for an opening statement.

Mr. GOODLING. No opening statement.

Chairman KILDEE. If there is none, any opening statement from Mrs. Lowey?

Mrs. LOWEY. I'll reserve my statement, but I would just like to welcome David Kearns and tell you how very anxious we are to hear from you and to work with you. We know of your commit-

ment to upgrading our educational system, and I thank you for appearing before us this morning.

Mr. KEARNS. Thank you, Mrs. Lowey.

Chairman KILDEE. Mr. Roemer?

Mr. ROEMER. I, too, would like to join in welcoming Mr. Kearns before our committee. I am anxious to hear his ideas on a very, very important topic. We are all looking for educational opportunity and enhancing that educational opportunity for all American children. We need to see some changes in the current system and we need to see some reforms in the current system, and we need to be creative about how we go about implementing these reforms.

I'm anxious to hear, after reading your testimony, some of your ideas in this regard. Given your background, I would also like to talk about the ideas on technology and how it can be used in our schools to enhance learning. I recently had the opportunity to see some interactive television technology and what that can do in our classrooms, so I'm anxious to listen to you and to listen to our distinguished panelists.

Thank you, Mr. Chairman.

Chairman KILDEE. Thank you.

Mr. Pastor?

Mr. PASTOR. Just to say good morning and welcome, and I look forward to hearing your testimony.

Mr. KEARNS. Good morning, sir.

Chairman KILDEE. Thank you very much. We have as our first witness a great friend of education who was a great friend of education before he had his present position. He worked very closely with it. It's a labor of love, and we appreciate that.

Sometimes, David, we can solve more in a 15-minute meeting out in the anteroom than we can here. I appreciated our earlier meeting here today with Mrs. Lowey and myself. I thought that was very, very helpful.

You come to us with great credentials: your love for education, your knowledge of the business world, having been former CEO of Xerox Corporation. We welcome your testimony here this morning. You may proceed in any fashion you wish.

STATEMENT OF HON. DAVID T. KEARNS, DEPUTY SECRETARY OF EDUCATION, U.S. DEPARTMENT OF EDUCATION, WASHINGTON, DC

Mr. KEARNS. Thank you very much, Mr. Chairman.

I really enjoyed working on the council with you and Bill Goodling. I learned a lot. As I say to a lot of people, I'm a sponge. I'm learning so much every day. It is a labor of love, and I do care about education. I care about the country, as all of you do.

I'll try to be brief. You've had an opportunity to read the testimony that has been submitted so that we have ample time for the questions. I am here today to express the Department's endorsement of the recommendations of the National Council on Educational Standards and Testing, and I appreciate the opportunity.

The council's report reflects 6 months of deliberation and enjoys broad support from a very diverse group, often a group that had been at odds on a lot of issues on past occasions. We concur with

the council's determination that real educational opportunity for all American children requires high national standards that are world-class and assessments that are based on those standards that are comparable, that are reliable, they are valid, and of course they must be fair.

World-class standards and high expectations for all children are the only way, I believe, to ensure excellence and equity in education. American children today are not achieving what they could be, and internationally test after test tells us that American students are doing poorly particularly in mathematics and science relative to the rest of the industrial world.

The NAEP data tells the same story. American achievement in science, math, reading, and writing has improved little or not at all over the last 20 years, which is certainly not good preparation as we move into the next century. The children with the most to gain from education, and that's the economically disadvantaged, are being tracked out of educational opportunity. If any group needs high standards, in my judgment, it is that group that deserves high standards most of all.

Now, some people believe that higher standards will adversely affect the disadvantaged youth and benefit only the brighter students, and we disagree with that. As we chatted earlier, I believe high standards are required for everyone, but different strategies will have to be applied in different areas so that all groups can meet those very high standards.

High standards and assessments, assessments that are tied to the high standards, will help solve rather than aggravate the problems of equity. If we continue to impose mediocre expectations on our young students, then we will continue to get mediocre results.

Higher expectations will inspire all students, including the disadvantaged, towards greater achievement. I'm convinced that over the last 25 years in our Nation not just in education, but we have continued to lower the standards of what this great Nation, in fact, can achieve.

If we continue to impose the mediocre expectations, then I know we will not do what this Nation can do. We cannot be complacent about this situation. As I wrote 3 or 4 years ago, some people think in typical parents' hyperbole, but I believe this, in fact, is a survival issue for our Nation; but it is clearly a problem that we can solve.

The vitality of our economy and our democracy depend upon a well-educated citizenry. Yet, test scores and reports from college professors and employers all indicate that we are graduating students unprepared for challenging jobs, for higher education, or for competent, knowledgeable participation in our democracy.

We all know that Americans like to be first, and by golly, we can be! In fact, I'm offended by the concept and discussion that we ought to have schools that are as good as the Japanese. We ought to have the best school system in the whole world, that is uniquely American, and we can do that. We know that high standards and high expectations for all children are the critical ingredient for developing a world-class education system.

When we look closely at what nations scoring highest on the international assessments are doing, we find one constant: a de-

manding curriculum based on high standards and assessments that are tied to those standards. Real improvement in education will depend on the States.

The Department will be supporting the States' efforts to change their whole education system based on world-class standards: National standards in the core subjects, State K through 12 curriculum frameworks and teacher enhancement based on those standards, assessments that are valid, reliable, fair, and comparable.

I would like to be very clear: we do not support a national curriculum or a single national test. The Federal role is, rather, to support the States, their best efforts as they move towards a voluntary system.

Separately, we know that some have expressed concern about school and system's performance standards. We support such standards. We simply think that they should be developed by the States and not mandated by the Federal Government.

Two basic reasons for this: First, we are encouraging flexibility in school-based management, not Federal mandates for school system management; and, second, we do not want to do anything that would discourage innovations in our communities across the country.

We recognize and appreciate your need to respond to the council's report, particularly to the recommendations regarding the coordinating structure. Whatever you choose to do, we believe you should start with that consensus. This is truly a unique opportunity to eliminate mediocrity as a benchmark for achievement in our Nation's schools, and we urge support of the Congress for the council's recommendations.

Mr. Chairman, I would be glad to answer any questions that you have.

[The prepared statement of Hon. David T. Kearns follows:]

Deputy Secretary David T. Kearns

Chairman Kildee, Members of the Subcommittee on Elementary and Secondary Education, I am pleased to have the opportunity today to express the Department of Education's endorsement of the recommendations of the Congressionally-appointed National Council on Education Standards and Testing, Raising Standards for American Education. Thank you, Mr. Chairman and Mr. Goodling, for your efforts on the Council that produced this historic report. It was a privilege to work with you. As a member of the Council, I can attest to how thoughtfully and carefully these recommendations were developed. The Council's consensus is a tribute to the hard work, commitment, and nonpartisan spirit of all of its members.

The report and recommendations of the National Council on Education Standards and Testing mark an important transition in American education. The Council has determined that real educational opportunity for all American children calls for national standards that are world-class and assessments based on those standards that are comparable, reliable, valid and fair.

In 1989, the President and the Nation's Governors met in Charlottesville and arrived at six National Education Goals. In announcing AMERICA 2000 the President sought to move the Nation forward in realizing the objectives of that historic partnership. AMERICA 2000 calls for world-class standards that will define

what our students need to know and be able to do, and for a system of voluntary American Achievement Tests that will measure progress against those standards. A recent poll shows that the public strongly supports world-class standards and assessments. We find the Council's report an affirmation of that national, public call for world-class standards and assessments.

Let me be absolutely clear: the Department of Education does not support a national curriculum. Neither do we support a single national test. The Federal role in acting upon the Council's recommendations will be limited to providing support and leadership for State and local efforts to implement the standards. I cannot overemphasize the point that the central and critical role in implementing the Council's recommendations will be played by the States.

Why do we need standards?

World-class standards ensure both excellence and equity in education. We need to improve the performance of all children: at present few of even our best students are getting a world-class education. Moreover, standards will help guarantee real educational opportunity for disadvantaged students.

Several weeks ago the Department of Education and the National Science Foundation (NSF) released the results of the second International Assessment of Educational Progress, which compared

the math and science performance of students in different nations. American students on the whole did very poorly. The Department of Education's National Assessment of Educational Progress (NAEP) data tell the same story. Over the last twenty years the achievement of American students in science, mathematics, reading, and writing has improved little or not at all--hardly good preparation for the competitive rigors of the 21st century.

Every student needs an excellent education; but the students with the most to gain from a good education are our nation's disadvantaged children. Research shows, however, that in our present system the most disadvantaged children--children of low-income families, children of parents with less than a high-school education, children of certain race-ethnic groups, children with limited English skills--are usually relegated to education "tracks" that do not prepare them adequately for employment after graduation or for higher education. There are, of course, schools and programs and individual stories that are exceptions--but they attract attention because they are just that: exceptional.

We know what second-rate achievement does to future economic competitiveness and that all our children must be prepared to compete with their peers throughout the world. Yet, critical as economic competitiveness is, the other contributions that

education can make to the life of our nation are no less important. Good education nourishes and enriches our humanity; it equips us both to sustain and to understand our democratic institutions; it enables us to speak to one another, to tolerate and appreciate our differences, and to see what it is that most connects us as a nation and as human beings. Yet our present system provides such an education only to an elite.

National Standards

The Department of Education already is supporting the development by major independent organizations of content standards for what students should know and be able to do in the core subjects. The mathematics standards developed by the National Council of Teachers of Mathematics (NCTM) provide an excellent example of how standards can catalyze change. The NCTM standards, which the Department of Education endorsed, were developed in consultation with the Nation's math teachers. They already have led schools, districts, and States to begin realigning their curricula and education practices with them. We are using that approach as a model for other efforts to develop standards.

Through the National Academy of Sciences (NAS), we are now supporting the development of world-class national standards in science. In cooperation with the National Endowment for the Humanities, we are also supporting the development of history standards by the National Center for History in the Schools based

at the University of California at Los Angeles. Both the National Academy of Sciences and the National Center for History in the Schools will develop a credible consensus process, broadly engaging teachers, scholars, and specialists in their fields in order to determine what American students should know and be able to do in these core subjects.

The Department also intends to support the development through the same process of national standards for English, geography and other subjects. As with the NCTM standards, these standards will be available for the States to draw upon as they plan their strategies for system-wide change.

In addition, the Department of Education's Office of Educational Research and Improvement (OERI) will be supporting international research on standards, particularly research that looks at other industrialized nations' standards and at the way standards guide education in these countries. We expect that this research will assist the States in developing a concrete understanding of what world-class standards and standards-based, system-wide change are all about. For example, the President has called for an Asian/Pacific Education Ministerial to be hosted by the United States to discuss world-class standards, particularly for primary and secondary education. This meeting will focus on how nations develop, implement, and assess achievement of those standards.

Supporting State Development of K-12 Curriculum Frameworks

We believe that equitable delivery of education requires each State to have clearly defined K-12 curriculum frameworks in each subject; and the Department of Education will support the States in developing their own K-12 curriculum frameworks. These State frameworks will establish well-planned, non-repetitive curricula in each subject to ensure that all children have the opportunity to study challenging subject material in every grade, K-12.

Beginning in FY 1992, OERI will be making grants through the Eisenhower National Program to assist States in the development of K-12 mathematics and science curriculum frameworks. And in FY 1993 OERI will hold a grant competition to support State efforts tied to standards to improve teaching and learning in other subjects--including history, geography, English, and foreign languages.

To improve teachers' subject matter knowledge in mathematics and science and to improve math and science teaching, in FY 1993 the Department will begin supporting an in-service professional development program in competitively selected school districts through the Eisenhower National Program. The Department also plans to work with the Organization for Economic Cooperation and Development on a conference in June of 1992 that will discuss research on world-class standards and assessments.

Separately, we know that some have expressed concern about school and system performance standards. We agree with the Council that those standards are important; we simply think that they should be developed by the States. There are two basic reasons for this: first, we recognize that a system already top-heavy with administration hardly needs Federal oversight of school system management; second, we do not want to do anything to discourage innovation.

Assessments

Assessment was the most complex and difficult issue that the National Council on Education Standards and Testing confronted. Much of the testimony you have received in the preceding hearings has addressed the complicated issues that surround testing.

Let me repeat: the Department will not support or encourage a single national test. We support the Council's call for continued use and refinement of the National Assessment of Educational Progress (NAEP) as the benchmark for how well our nation as a whole is meeting the goals, and as a benchmark whereby States can gauge how well their children are meeting the goals. For assessing individual achievement, we will be supporting the development of multiple new assessments by States or groups of States. As the Council's report emphasizes, it is critically important that these assessments be valid, reliable, and fair. The different assessments should also provide

comparable indications of whether students are meeting world-class standards, whichever test they take.

We do not suggest that developing a fair, reliable, valid, and comparable system of assessments will be quick or easy. What we are saying is that we must get the process underway. If we had insisted that the first copying machine be perfect, we would still be using carbon paper--on manual typewriters. Developing assessments will be an ongoing process, and we need to begin.

Make no mistake: students are being tested as we speak--and with tests that we know narrow the curriculum. We already have assessments in place that have proven inadequate to our needs and expectations and detrimental to our children's and our nation's futures. It is time to replace them with something better.

We are not recommending simply fixing the tests we have been using. We agree with the Council that we must instead build a new system of assessments, working (as are California and other states) to develop assessments that measure more fully than the traditional multiple-choice test what students know and are able to do in each subject. These assessments will not spring into being full-blown and flawless but rather will evolve and improve over time. That is simply the nature of good research and development.

We do not support attaching "high stakes"--college admission and employment, for example--to the results of untried tests. But the failure to act is as much a decision as action itself--and the stakes of our inaction mount with every year that we continue to watch shallow tests dictate what our children study and what our publishers publish. While we do not support high stakes being attached to tests at this time, we do hope that in the future high achievement on valid, reliable, and fair tests will have consequences for students' futures. Students will then see the link between their aspirations for the future and enrolling in challenging academic courses and doing their best. Good tests will measure the achievement of high standards; and high standards demand good tests.

The Department will therefore be supporting research and development of assessments. For example, we plan to offer incentive grants for assessment development in 1992 through the Fund for Innovation in Education (FIE). Under FIE, the Department will hold a competition for State and local projects to develop approaches to assessing children's achievement in mathematics, reading, and writing in the fourth grade. Applicants will be invited to develop their own models for such assessments in collaboration with classroom teachers, subject matter specialists, and experts in tests and measurement.

BEST COPY AVAILABLE

These State and local projects will serve as living laboratories for research on assessments in FY 1993. The projects will provide a context for research on the reliability, validity, and fairness of various assessment instruments and testing conditions, as well as on the equating and linking necessary for making differing systems of assessment comparable.

Standards and Equity: Educating the Disadvantaged

The Council's report notes that some observers have expressed concern that at-risk students will be harmed by the creation of world-class standards, curriculum frameworks, and assessments based on these standards. Some say that establishing national standards simply raises the hurdles and dooms more of these students to failure. We disagree.

All students deserve to be taught the best that can be known. All of our students should have the opportunity to learn those things needed for further education and for the world of work, and many of those with the least such opportunity are disadvantaged children. Our research has shown that high expectations are a necessary ingredient of educational improvement, particularly for disadvantaged students who presently suffer the most severe consequences of low expectations. Setting standards and developing meaningful curricula and assessments will help identify the kinds of instructional assistance disadvantaged students need and will

help improve the instruction they receive.

The Department plans to assist the States by drawing upon the existing range of work it already supports for improving the education of the disadvantaged. For example, OERI supports four university-based national educational research and development centers devoted exclusively to work on issues related to at-risk students; and projects in every one of the national research and development centers that OERI supports address ways to improve the educational performance of at-risk students. The Department of Education plans to conduct new research on effective teaching methods and curricula for at-risk students, and to enhance efforts to make the results of that work available and accessible to all teachers and parents.

The Department, through the Congressionally mandated National Assessment of Chapter 1, is also looking carefully at the earlier standards that were set to gauge the progress of disadvantaged children in Chapter 1. The new Chapter 1 Testing and Assessment Panel will consider these Chapter 1 standards and the way that the testing required by Chapter 1 could be aligned with the kinds of assessment needed for the Nation, as well as other recommendations for change. Let me reiterate that we believe all children--whatever their background--are best served by high standards.

Conclusion

The Department strongly endorses the recommendations of the National Council on Education Standards and Testing and intends to assist the States in bringing about the changes necessary to prepare all children for competent, productive, thoughtful lives in the 21st century. We urge the support of Congress for high standards for all youth and for system-wide change in each State to help students reach the standards: the development of State curriculum frameworks, new teacher education and certification requirements, and assessments that are valid, reliable, and fair. We support the establishment of a coordinating council as described in the Council's report, and we also urge the Subcommittee's support for a House measure similar to that passed in the Senate for this purpose.

Mr. Chairman, that concludes my prepared remarks. I will be glad to answer any questions that you or the members of the Subcommittee may have.

Chairman KILDEE. Thank you, Mr. Secretary.

Has the administration provided grants to any subject area groups to develop national standards? Could you tell us something about that? Were these grants competitive that were——

Mr. KEARNS. Mr. Chairman, the Department has provided some dollars to help develop standards through the National Academy of Science, through the National Endowment for the Humanities, and have been supportive of the National Council of Teachers of Mathematics; so, in the subjects of math, in the sciences, and in history.

Regarding the question on competitive, in the area an issue has come up to do with history standards. The dollars went from the Department to the National Endowment for the Humanities, and they made an addendum to a contract that they already had with the Center for History in California. It did go to their board and did have peer review.

Mr. KILDEE. David, some people indicate that standards are the equivalent of curriculum. Could you comment and see what distinctions you make over in your shop between standards and curriculum?

Mr. KEARNS. Well, Mr. Chairman, I do not think that that is right. The setting of high standards of what children should know and are able to do is then a first step. The States, working with local communities and, hopefully, groupings of States, will then develop assessments and curriculum frameworks for what is actually taught. We believe that that should be developed at the State level and that they should be developed with a broad range of inputs as the standards should be developed.

If you look at the National Council of Teachers of Mathematics and the process that they went through in developing those standards, I was impressed with that, that it was a broad input. They went to a large number of groups, got that input, putting those standards in place, and now the curriculum frameworks will be developed at the State level. I think that there really is a difference, and I strongly believe that we should not have a national curriculum, but I also think we should have high national standards.

Chairman KILDEE. I ask the question because I think, during the course of enactment of legislation to implement the commission's findings that people will ask that question. I think that both we in the Congress and you in the Executive Branch have to be prepared to respond that we are not in the process of establishing a national curriculum; that the local school boards and the State school boards have the wisdom to decide how to reach these goals, and the various methods that may be used to reach these goals. I think it's important that we respond to that because those questions will be asked during the course of this hearing.

Mr. KEARNS. I do agree with that. I think it's very important, because I think that the country basically shouldn't have a national curriculum. I don't think they would buy it if we tried to push it at them, in any event. We do want to be supportive, from a department standpoint, of the States as they develop their curriculum frameworks.

Mr. KILDEE. One of the questions that came up—and I'm going to defer then to the other members—is how we improve education with testing rather than just gathering data. We've given a lot of

tests, we can gather a lot of data, a lot of information. I can predict in certain schools what the test results will be, fairly accurately.

Suppose we, as we discussed earlier before the forum started, that in, say, the fourth grade we were to have standards in math and standards in reading and standards in the various subjects, and we then were to test and find out that in those schools, certain schools, students were not reaching those standards.

Would it not be appropriate, then, that with a combination of Federal, State, and local funds that some type of programs be developed for those schools to help those students come up to the standards which they had failed to come to up to as determined by the assessment?

Mr. KEARNS. Well, I agree with that. I think that's really the purpose of assessment is to understand how our youngsters are doing and what they know and are able to do, and then to use those assessments to develop specific programs. As I say, we need high standards. We need to assess how people are progressing against those, and those should assist in developing particular strategies by State, by local community, and, in fact, by school.

Mr. KILDEE. So standards, assessment, and then intervention or remediation?

Mr. KEARNS. Yes, sir.

Mr. KILDEE. If the assessments indicate that they are not reaching the standards, you feel that would be a good consequence of it?

Mr. KEARNS. I do. As we also talked before, I think that is the appropriate sequence. I don't think that we should wait to develop and to fund work that's going on, on new types of assessment. I think everyone agreed, and as we all saw in the testimony that took place during the council's deliberation, that there is a lot of new ideas and thoughts out there. I think that the government should be supportive of performance assessment, portfolio assessments, and some of the new ideas that are taking place and, I think, are currently being developed by some of the testing groups.

Mr. KILDEE. Thank you very much, Mr. Kearns.

Mr. Gunderson?

Mr. GUNDERSON. Thank you very much, Mr. Chairman. Thank you, Secretary Kearns.

I'm going to assume that on page 6 of your testimony regarding the Eisenhower National Program and efforts for in-service professional development—I'll let you find that at the bottom of page 6.

Mr. KEARNS. I've got it.

Mr. GUNDERSON. I'm going to assume that that was your idea as a result of your visit to western Wisconsin and your conversation with UW La Crosse and their desire to establish a program to assist rural teachers of math and science in improving their abilities; is that correct?

Mr. KEARNS. Of course, and certainly other trips and so forth. Steve, I did enjoy being out there and we did have that conversation about that.

Mr. GUNDERSON. Actually, my question is, to what degree can you accelerate from fiscal year 1993 even into fiscal year 1992 the opportunity to fund those kinds of initiatives? I mean, to be honest about it, we've not yet found in the Federal Government where we

can get the funding for those programs like UW La Crosse, to make happen.

Mr. KEARNS. Well, I wouldn't be sanguine about our ability to do things in 1992. We have a relatively small discretionary budget to fund those competitive sorts of things. Particularly in the math and science area, it's coming out of the fix-it strategy, we are trying to move as much as we can into that direction and then continue to do that.

We signed an MOU with the National Science Foundation to try to also coordinate the efforts in teacher education between NSF and the Department. But we are well into 1992 at this point in time.

Mr. GUNDERSON. Well, I don't want to keep pushing the same issue, but I have to tell you we have 2 weeks in which if we can pass a Neighborhood Schools Act, a piece of legislation, you will get a hundred million more dollars in fiscal year 1992 to use for this program than you would have otherwise.

Mr. KEARNS. Yes, we did have that conversation before, and, yes, you are continuing to push it. I would just say to that, as you work your way through, however the legislation that you are currently working on comes out, that it is important that however those dollars be spent, that all of us make sure that the package is put together in a way that it really goes towards real reform and new thoughts and ideas.

Mr. GUNDERSON. You indicate in your statement, and frankly you use as an example for national standards the National Council on Teachers of Mathematics. I think that part of the great debate about national standards is a misunderstanding or a lack of clarification as to exactly what do we mean by the adoption of national standards.

Now, if I understand correctly what you are saying in your statement, you are saying that groups like the National Council on Teachers of Mathematics will adopt standards. Does that mean that the Department also adopts standards, or that there will be various groups out there adopting standards which, hopefully, people will try to reach?

Mr. KEARNS. Well, I can't predict how it will all come out, but the National Council on Teachers of Mathematics develops those, and they get a broad input that you would expect that the States, you would hope that the States would accept those standards.

Mr. GUNDERSON. Okay. What I'm trying to get at is, okay, we've got the National Council on Teachers of Math; they've got their set of standards. Let's assume the National Academy of Sciences also develops a set of standards, and let's assume that some research group at Harvard University establishes their set of standards. We've got three different sets of standards for math education in this country.

Does the Department take a little of each and adopt through regulation a set of national standards? Do you allow, then, that menu of options out there for anybody and everybody to choose which ones of those they want to achieve? Here is where I think the lack of clarification and a misunderstanding or lack of understanding exists in the education community and why this issue is even controversial. Take me from the National Council of Teachers of Math

and where we go from that in terms of achieving your department's goals on standards.

Mr. KEARNS. First of all, it would not be—I think you have to go back and you take a sequence that Dale went through just a moment ago, and I'll just take you back one further. It's the six national goals, which I believe there is a consensus on, that we should be driving towards those. Then you come and then you start yourself through a sequence. In the standards, the thought process is that the standards would be developed, that they would be voluntary, that people would accept those and the thought process of the States——

Mr. GUNDERSON. I hate to interrupt you, but who develops the standards?

Mr. KEARNS. Well, right now the National Council of Teachers of Mathematic is, in fact, developing standards for mathematics.

Mr. GUNDERSON. Have they been authorized by the Department to do that?

Mr. KEARNS. No.

Mr. GUNDERSON. They are doing it on their own?

Mr. KEARNS. Well, there has been funding to support their activities, but it's not an authorizing body. Then if you go back and if you look at the panel that Dale and I were on—I was going to say INCEST, the council had just went through—what it suggested is after the goals panel gets reconstituted, that there then would be a council that would put an imprint on it. It's not really an authorizing body, I think I would call it an advisory body, that would put an imprint on that.

The National Academy of Sciences has been funded to start developing standards in the sciences. And those groups are working on a very broad basis. Then move through history and next would become geography. But it would not be—the Department of Education wouldn't authorize that as the Federal standard.

Mr. GUNDERSON. Okay. What happens if MIT also develops a set of standards in math?

Mr. KEARNS. Those standards would then—the idea of the council would be is that if those were good standards and so forth, they could be authorized as high standards and the States would then either accept those or not accept them.

Mr. GUNDERSON. There is not going to be the imprint of the Federal Government on one set of standards as being "the standards?"

Mr. KEARNS. I think that may, in fact, be the way it would come out, but I don't think anybody has said that there only could be one standard if there was another body that developed another standard and it made sense. As you take the whole thing through the track, the one thing that you do need—the assessments will be voluntary, and the States will do their own curriculum framework.

We have said that we thought there needs to be an ability of comparability when you get to assessments so that what is happening in one State, that there be a mechanism that would compare results not for individuals but for the system itself.

Chairman KILDEE. Would the gentleman from Wisconsin yield a moment on that point?

Mr. GUNDERSON. Sure.

Chairman KILDEE. We're still working our way through the structure that's going to be put together, and I think your question was very relevant to that. One structure could be that we would have this national entity which we kept talking about during our council meetings. We don't know what——

Mr. KEARNS. Whatever it is.

Chairman KILDEE. We don't know what that will be. Then the National Teachers of Mathematics may develop standards, and then this national entity may give the "Good Housekeeping Seal of Approval" to those standards.

Mr. KEARNS. Right.

Chairman KILDEE. Then the local, the States, can voluntarily agree to accept those standards. Is that fairly well the structure?

Mr. KEARNS. That's my understanding.

Chairman KILDEE. Whatever the entity might be, it would give the "Good Housekeeping Seal of Approval" to, say, the National Teachers of Mathematics' standards. Then it is still up to the States to voluntarily accept those standards.

Mr. KEARNS. Yes, sir.

Chairman KILDEE. Right.

Mr. GUNDERSON. Okay. Mr. Chairman, then, let me ask you a question.

Chairman KILDEE. Okay.

Mr. GUNDERSON. Only because I'm trying to get this clarified. I think there is a great misunderstanding in this area. Will we in the Congress authorize an entity that will then determine whose set of standards are the national standards?

Chairman KILDEE. We will probably authorize an entity who will give a seal of approval to standards. Possibly two groups could have standards that would get the seal of approval. It could be maybe one, maybe two, maybe more.

But then I would think, I am not going to guess the number, but then the States would voluntarily say: "These standards that have been set, say, by the NTM, given the seal of approval by this entity, we now accept as our standards within this State."

Mr. GUNDERSON. Okay. That's helpful.

Chairman KILDEE. We're still working through this, you know, but I think we're following some of the recommendations that were deliberately left a little loose for the Congress to act upon. But I think that is a possible scenario.

Mr. KEARNS. I agree with that.

Chairman KILDEE. Thank you.

Mr. GUNDERSON. Well, I hope we develop this "Good Housekeeping Seal of Approval," and I know I've created more questions on the other side of the aisle by looking at my colleagues faces so, Mr. Chairman, I will yield back my time.

Mr. KILDEE. Well, no, you and I approach things structurally, and I think that's very good. I think we have to look at that. Very good.

Mrs. Lowey?

Mrs. LOWEY. Thank you, Mr. Chairman. And I want to thank you, Mr. Kearns, again for your testimony.

I would like to share with you a dilemma which I am currently struggling with, and perhaps you can provide some guidance for

me. You discussed with the chairman your view of this issue in terms of standards, assessment, and then remediation and that it really has to go together if we're going to make a major change. Clearly, your approach to this whole problem of educating our youth for the future is a major sea change, and it's a major challenge.

Now, what concerns me, currently in New York State the Federal Government provides 4 percent of the education dollars. As I go back to my community and so many of us do, we are constantly facing taxpayer groups that are voting to cut budgets. The States have had to cut back on their dollars for education. That's in spite of the general commitment throughout the country on every level that education is the key and education is the key to our competitiveness and the key to the future, so that they are feeling squeezed, and they don't know how they are going to pay for their current budget.

My concern is we all agree on standards. I assume we are going to work towards that end, and then we put in place assessment procedures. My problem is, are we raising hopes and how are we going to get to the remediation? I look at England, for example, that provides, what is it, between 80 or 85 percent of the dollars for education on a national level.

We know, for example, that Head Start works. We know from the studies of Carol Gilligan and others that the girls at the age of 13 and 14 are turned off of math and science, and we have to inject resources and support services at that age in order to really get these young girls into the fields of math and science.

We know today, and in many parts of my district we know for sure, that all the problems of our society converge on our educational system. We know that we have to link up resources, health services, parent training, all the other social services that exist. We know we have to link them up with our school system. We know we have to get drugs out of our school system.

My concern, Mr. Kearns, is having another study, another set of standards, having them sit on the shelf without any real plan as to how we're going to inject the resources and build up the support to get those resources into our community to really doing the job. The States are not going to do it; they can't. The locals are suffering; they can't.

How do you view the Federal role in the remediation area, because I support everything else, how do you view the Federal role? Should the Federal Government assume a greater responsibility for the resources in providing the success that we hope to achieve from standards and assessment?

Mr. KEARNS. I can hardly wait to hear what I'm going to say next. It's a broad range of points that you've made, Mrs. Lowey. Let me start by saying I think one thing we have to watch is, is that we don't catch ourselves as a Nation or at the community level of what I call a serial approach to solving problems, that you can't solve problem two until you get one done.

What I'm concerned about that is that we have to solve the drug and the violence problem. There isn't any question about it. It's getting in the way of educating our youth. No question about it at all.

The integration of social services and education is absolutely imperative. The successful schools that I visit have done that. We have to increase substantially the knowledge base of our teachers that are out there, particularly in math and science. There is a broad range of issues that have to be solved.

We are also going through a tough economic time right now where the sales revenues at the State levels have made it very tough, but that, too, will pass. The Nation in fact has at the local level, the State level, and the Federal Government level over the last 12 or 13 years, has spent a lot, has increased the amount of money on elementary and secondary education by a lot.

Now, specifically to your question about the dollars. Head Start is an effective program. We support increased funding in that area. Second, there are Federal programs, and we have an opportunity as we go through reauthorization for elementary and secondary school programs, which Chapter 1 is included, and we chatted about that this morning, for next year.

In looking at more flexible spending of those dollars as they relate to State and community strategies, and the idea of flexibility is currently being worked on so that the dollars that are spent could be used more effectively. I don't know what the exact right amount of money is to be spent on education. I do know that we spend—it's tremendously varied, and outcomes are not necessarily tied to the amount of dollars that are spent.

Mrs. LOWEY. Right.

Mr. KEARNS. I also know in looking at school budgets across the country—not necessarily just in big cities but in others—that we spend, in my opinion, an inordinate amount of money not at the schoolhouse where the children and the teachers are.

I do think that there are Federal and State programs that can help with these strategies, but I do think that the strategies are going to be developed at the community level and that we may need State legislation to make it possible also for the communities to develop what they need to do.

We are really the only industrialized Nation that does not have a national education system. It's interesting to have this other discussion and then move in this direction. I don't know exactly what the right percentage is that the Federal Government should spend.

I do know this, that I don't think that it is politically doable, nor do I personally agree that we should do away with the State and the local system and go to a basically Federal Government-supported education system across the country. We currently spend about 8 percent, on average. It's apparently less of course in New York State.

Mrs. LOWEY. Right.

Mr. KEARNS. That's a tough question, but I do believe that the communities can develop strategies. I am also convinced that when a strategy is out there—and I think high standards and assessments will help us with this—you can find it in communities across the country. When the local community understood the strategy, they will fund it. I do believe in the current state in this country that when they don't understand and they don't think that there is a strategy, that they won't vote for more money into the system.

There are some excellent examples of school superintendents that have built strategies. One was John Murphy in Prince George's County, right here where the community ended up understanding the strategy and 2 years into it, it got a 15 percent increase in taxes in the school—which is almost unheard of.

Mrs. LOWEY. Well, let me say this. I do hope that we can continue some informal discussions as we had before because I think it's very helpful. I, too, agree that I'm not sure exactly what the appropriate breakdown is between Federal and State. I do know that there are rebellions all over, not only my district but in other districts.

I do feel that too much is spent on administration; it doesn't get right down to the classroom both on the State and local level. But I also feel that there are certain givens that everyone accepts be it Head Start or resources for 13- and 14-year-olds.

As we begin to discuss the reauthorization of the Elementary and Secondary Act, I think this discussion should be front and center, and if we have specific goals in mind, perhaps we have to rethink those percentages and perhaps the Federal Government does have to take an even greater responsibility for infusing essential resources into our local school systems but——

Mr. KEARNS. Well, we look forward to working with you.

Mrs. LOWEY. Thank you.

Mr. KEARNS. We do all have an opportunity as the hearings and the input are just beginning now in elementary and secondary reauthorization, starting around the country, and we will have an opportunity to work together over the next year, and I look forward to that.

Mrs. LOWEY. I look forward to it. Thank you.

Mr. KEARNS. Thank you, Mrs. Lowey.

Chairman KILDEE. Thank you.

Mr. Roemer?

Mr. ROEMER. Mr. Secretary, let me ask you. You just said, I think very articulately, that we have a host of problems in education. I will soon be going out the door to meet with 50 children from my home district in North Liberty, Indiana, talking to them about what we've discussed here in this committee, as well as what we do as congresspeople on Capitol Hill.

It was interesting to note yesterday in The Washington Post the recent poll that said that 54 percent of our children do not think that they are going to be able to do better than their parents. That is not a lot of optimism on the part of our children. Therefore, education, I think—I agree with Mrs. Lowey—has got to be a higher priority in this country. We must see our businesses and our educators working together and our government being part of the solution, although money isn't the only solution in this dilemma.

Going back to the first step we take with at-risk children, you talk in your testimony about university-based national educational research and development initiatives for at-risk children. Could you explain that a little bit more and talk about how we get our universities working on at-risk children, working with curriculum—maybe not national testing and national curriculum, as you do not endorse it? But what do we do to get our universities involved in this problem?.

Mr. KEARNS. Let me separate it into two pieces. We are supporting, there are four centers—Franci is back here.

What are the four? Have you got them? They're digging.

It's in my fat book here—but that are regional centers around the country that are being funded to do work specifically for at-risk children. This is separate from the discussion that we were having before about supporting States in curriculum framework. That is specific work. I will dig out where those four places are.

Now, on the curriculum side—and I get in trouble with the university community from time to time about what I'm going to say next—is that if you look at other areas of competitiveness of the country, whether it be semiconductors, biotechnology, or optics, you will find a combination of the Federal Government, private industry, and our research universities working together to keep this country out front.

Frankly, I do not see the university community in education playing a leadership role in driving education reform across the Nation. That's a challenge for us and it's very important. There are obviously exceptions. I quote Ted Sizer and Hank Levin and Hal Stevenson from Michigan State and Stanford and Brown and different schools. They are some very good pieces of work, but they are not playing a leadership role.

We were talking about curriculum frameworks and standards. Our schools that are educating our future teachers and are often providing in-service training of teachers will have to reach for those same high standards, change their curriculum to ensure that all of our teachers have the appropriate training as well.

This is a massive job. What all of us are working on here and talk a lot about, about reform of a system—I've worked on a lot of tough things in my life, but the hardest thing I have ever been involved in is trying to get a university or college to change a curriculum. It is really tough, and I think that we are going to have to bring some amount of pressure to bear, and I think we need to put incentives into the system.

Hopefully, as we develop again down the road for teacher training that Federal Government dollars ought to go to teachers that then go to a school that's reaching for these new standards and has changed the curriculum. In other words, I think we might be able to put some incentives into the system.

Mr. ROEMER. Mr. Secretary, in addition to curriculum, what can universities do to work with these children in the first and second and third grades and as they first walk through the doors into kindergarten?

I just talked to a principal in my district the other day. His school is in a rural area, the largest town is about 20,000, so it's not an inner-city school. He says that about 35 percent of the kids that walk in the first grade are at risk of dropping out right away.

The locally devised, voluntary test of skills that we give to these kids in the first couple of grades, do the universities work on developing that? Do the teachers through in-service programs work on developing that? How do we assure that a meaningful test is developed for a place like La Porte, Indiana, and not only in Los Angeles, California?

Mr. KEARNS. Well, I'm not sure that I know the answer about how we do it at each school, but I would also tell you that much of the—that in many of the local communities there are universities that are working with the schools. I happened to be in Chicago last week, and De Paul University is working with a set of schools right there in the community. Steve and I met in Wisconsin, out in La Crosse, last summertime with similar things. So the schools can be involved and so forth in doing that.

Let me make one comment about a program that I think would both give you hope, but also says there is nothing basically wrong with our children in this country. I made this comment earlier, I think, outside. Hal Stevenson at Michigan State University has done a number of studies. Basically, it says that our children from out poorest communities come to school and they are not behind the East Asians, in other words their basic knowledge. This is from our poorest communities.

Second point. Hal or Hank Levin at Stanford University has developed a program he calls accelerated learning. It is now into about 140 schools around the country: big cities, rural settings, and others. Basically, what he has done is taken programs and curriculum that were developed for gifted children, has given them to our poorest children from our poorest neighborhoods, and they can do that work. In about 2 years, they are at grade level and start to move ahead.

Now, what that basically says is, is that what you need is demanding curriculum. What we do often now is that we track kids out early because somebody determines that they can't learn. We should basically believe that all children can learn. That's the underpinning of the six national goals. Then all programs should be developed on that basis. Some children clearly need more attention, as we talked about before, than others. I think our universities can be helpful, but I do not think they are playing a leadership role today. Johns Hopkins and Boston University are two of the centers that we're working with, and there are two others.

Mr. ROEMER. Can you get those back to me in writing?

Mr. KEARNS. Yes, sir.

Mr. ROEMER. Finally, Mr. Secretary, I appreciate your patience here. We hear and I have heard both on my Science, Space, and Technology Committee and the Education Committee, different CEOs testify that businesses are spending between $30 billion and $35 billion a year reeducating students once they get out of high school.

As a former CEO of Xerox, what would you tell business leaders that want to take a role—what specifically should our business leaders be doing in our communities to get involved in this education reform movement?

Mr. KEARNS. I think you said million, it's billion; and maybe that's what you meant.

Mr. ROEMER. I meant billion.

Mr. KEARNS. Yeah, it's some place on the high side of $25 billion, and getting larger. I think business has an important role to play, and it's a broad range role. First of all, business leaders in the communities ought to support the educators and the politicians that are out in front for real education reform. People like myself find

the political environment hostile. We don't like it and tend to shy away from it.

The changing education that is in this country is a political process. It will take legislative change at the State level. Business leaders need to get involved in that in the business round table. The larger companies have been involved in that with their nine-point program to work at the State level for legislative reform. They need to work at the local level.

Again, you can go into different cities and different communities and you see different companies that do a broad range of things all the way from what I would call the negatives, which are what I call "feel good" programs, that make the company feel good but don't have much to do with systemic change or reform, all the way to real heavy involvement at the local level with their dollars and with their people. Often, the human resources that a company can loan are often more important than dollars.

Xerox has had two to three people on loan to the Rochester New York School System for the last 3 years on total quality management, the process that we put in place to turn our company around. I don't think it's appropriate for the business leaders to be involved in curriculum develop. Some of my associates disagree about that, but I don't think pedagogy and curriculum development is what they should be involved in. But they need to be supportive and they need to be involved in community-based strategies.

Now, in addition to that, businesses across the country have been supportive of the New American Schools Development Corporation, which is a private nonprofit enterprise to look at a national level at reinventing the schools for the next century with all new—with new ideas, new learning environments. As I like to say, learning environments 0 through 20 years of age, and to bring together the best thinkers in America.

That process took place last fall, and the bids came in 2 weeks ago, and 686 different teams of people from 49 of the 50 States including: 246 different school districts; 46 different individual schools; 140, by the way, universities; 136 different community-based organizations with one from the elementary school principals, to the AFT, to community service organizations across the Nation.

There are probably 2,000 different schools that are actually involved in that process. About a third of those proposals included high-technology approaches as part of the thing: a third of them came from inner-city schools; 25 percent from rural, poor areas; about 25 percent from suburban; and the other was a combination of those. Business can get involved in a broad range of activities, but it is in their best interest to get involved and to play a major role.

Mr. ROEMER. Thank you, Mr. Secretary.

Chairman KILDEE. Mr. Secretary, just briefly. You made a point that the universities could be doing more, and I actually agree with you. I think that they should be doing more.

You mentioned Michigan State University, and I was glad to see you mention that. I have a son who is a sophomore there, and another son who is a senior at University of Michigan too. But Michigan State is doing something very, very interesting in Flint, Michi-

gan where Secretary Lamar Alexander visited with Mr. Oakum a few weeks ago. They are putting some of their future training people right in Gundry Elementary School to work with the students and retrain the teachers, and at the same time getting ideas from that experience, taking it back for the initial training of new teachers at Michigan State.

Mr. KEARNS. That's excellent.

Chairman KILDEE. That is really a remarkable program and, I think, something that could be replicated by other universities throughout the country.

Mr. Pastor?

Mr. PASTOR. Thank you, Mr. Chairman.

First of all, let me commend you. A number of times in this committee I've heard where we have music teachers teaching calculus and music teachers teaching biology. I was wondering if you're going to go beyond the content of math and science and provide in-service training for other curricula, like music?

Mr. KEARNS. Mr. Pastor, in math and science there is legislation that the Eisenhower money, that is particularly legislated in the math and science area, and it has to be used that way. The National Science Foundation has programs specifically in math and science. In addition to that, we are supporting the States in the development of curriculum frameworks and assessment that go beyond those subjects, which would include the arts and languages as well.

Mr. PASTOR. Thank you.

On page 3, you write: "Research shows that in our present system the most disadvantaged children are relegated to education 'tracks' that do not prepare them adequately for employment after graduation or higher education."

Now, does that research show that the result of tracking is because teachers have low expectations of the students, or are the tests biased or unfair, or are the tests valid and, in fact, the students belong there?

Mr. KEARNS. Well, probably all of the above.

Mr. PASTOR. What's that?

Mr. KEARNS. Probably all of the above. There are probably some inequities in the test. Basically, the point that I was trying to make there goes a little bit to what Mr. Roemer was talking about, is that very early on kids particularly from our poor neighborhoods often get tracked out of a demanding curriculum, and low standards get set for them very, very early on. Therefore, when the low standards are set for them or low expectations, they then, in fact, don't perform. It's not startling.

That's why I tend to—I'm not plugging Hank Levin, but I have visited a number of his schools, and he takes youngsters and gives them demanding curriculum developed for gifted children, and they are able, in fact, to perform when high expectations are set and a demanding curriculum is given.

That's really the point that I was making is, that we can't track out youngsters early because they come from a disadvantaged area. They should have high standards. Then this is the point that Dale and I went on before, is that then a strategy for those children must be developed.

Mr. PASTOR. Then the assumption I'm making is that these national standards that yet need to be developed will give higher expectations for students. But at the same time, I think what we need to do is to develop those national standards with higher expectations, that as you admit that some of these tests may be invalid or may be unfair, that at the same time we, in a parallel course, maybe test or evaluate the current assessment systems so that we can validate them, and make them fair. I would agree that maybe we ought to be doing both at the same time.

Mr. KEARNS. Well, we are really encouraging that. In this work that the Department had suggested that we support, there is some exciting work going around about completely new assessments and different ways to go at it. We have to make sure that those are valid, and we must make sure that they are fair. No question about that.

Mr. PASTOR. I have one last question. I think what we're trying to do is develop a process, that this body is going to develop a process, in which we will develop national standards.

Hopefully, there will be a group out there that will either give them the good housing standard, or whatever the hell it is. We need to find a group that will basically say these national standards are what we need to define an educated person. I think we're looking for a process.

The process that you have in your written testimony is that we would have national organizations like the national math teachers, the national English teachers, the national science teachers who would develop some national standards. Then the States would develop some curricula frameworks. All this is voluntary in terms of the State, as I understand you?

Mr. KEARNS. Yes, sir.

Mr. PASTER. If we're trying to achieve a national standard, doesn't this voluntary position of States kind of defeat that purpose?

Mr. KEARNS. Well, when you get attacked from both sides—on one side, we say we're trying to force the national curricula, which we're not; that's why we've gone the voluntary route. Then of course you can come at it from the other way; if it's so important, why don't we just go down and say everybody do the same thing.

We feel that the best way to get this going is on a voluntary basis; that with a group that could put an imprint on the high standards, that that will encourage the States to do that. And if the assessment process, voluntary, does have a comparability element, in other words so that we would understand how different States are doing as they are compared to each other, we think that that would work. Trying to force-feed one national standard or one national test is not the right thing to do. I certainly understand the debate. I'm not trying to have an argument with you.

Mr. PASTOR. I understand.

Mr. KEARNS. We have come down in the middle of this thing to drive it. Given the fact that we are getting a lot from both sides, means that we're about right.

Mr. PASTOR. But the purpose of this whole nationalization is to find a definition of what skills make up an educated American. If

we try to say, well, we will let some States participate voluntarily, I wonder whether we will get to that national definition?

Mr. KEARNS. Well, I think we will. As all of you know probably better than I do, this is a big and an extraordinarily diverse Nation, and I think for somebody here to put an imprint that says this is exactly the way it will, and it will have to be done by every State and every community by some date, I think we would be having discussions here for a long time, and I don't think the American public would end up buying it.

I think, I feel the voluntary approach is the right way to go, and I think that when we set those high standards I think States will find that they will want to join in and not opt out.

Mr. PASTOR. Governor Roemer proposed another way of doing the procedure, having the standards developed from the bottom and work up nationally so that then you would have a consensus nationally of what the States would want. Is that another process that you would support?

Mr. KEARNS. Well, in some ways that is what is happening, I think. The States are out there working today. California, Vermont, Kentucky, there are other places that are well underway with a lot of this work. I think what we will find is that States will cluster together and not try to reinvent the wheel in a lot of cases. If there was a way to encourage this, both by the Department supporting the States' work in the development and the suggestion that the council made that's before you that could assist that process and, I think, it would happen more quickly than otherwise.

Mr. PASTOR. Thank you very much.

Chairman KILDEE. Mr. Petri?

Mr. PETRI. Thank you.

Thank you for being here this morning, sir——

Mr. KEARNS. Good morning, sir.

Mr. PETRI. [continuing] and for your enormous contribution both in the private sector and now as a public servant to advancing the dialogue and increasing people's awareness of the importance of the education enterprise and what can be done to make it more successful.

I just have one area to ask about and that is, did the commission give any thought to, or do you have any ideas about what we can do, or what we should be doing on a voluntary basis where people don't meet the standard?

Would it be a good idea to have tests of teachers and of parents as well as of kids? It may be that it's a whole classroom or a whole school that's substandard, rather than just a normal curve distribution where some do naturally better than others, that the fault lies not in the children but in some other circumstance that might be worth correcting?

Mr. KEARNS. Yes, sir?

Mr. PETRI. What I'm asking is, if we do the standards, then what?

Mr. KEARNS. There was a lot of discussion about that.

I don't recall, Franci, that we talked about testing parents?

Ms. ALEXANDER. No.

Mr. KEARNS. We did talk about the other things. In fact, we talked about systems performance. In other words, how well is the

system performing. I think most of us agree that we don't start on the basis that there is something wrong with the children. It is the leadership's responsibility to provide the appropriate education system so all the children can learn. We have been talking about that this morning, the idea of standards. Dale and I were talking about this before, the idea to have standards and then assessment.

Then that gives you the knowledge about schools or systems of schools that are not performing and that either remediation—a word that the chairman likes to use—or a strategy needs to be developed for that community for particular problems that either the system has or either particular problems that the students themselves might have because of their backgrounds or the lack of opportunity or parental support or whatever the activity.

There was a lot of discussion about that, and I think that it is an important piece of this. The reason for high standards is to know where we need to go, and the assessments is to measure how we're doing against it. Then you have to use those.

If all we do is bash ourselves and say we're not as good as the East Asians, that's not really very useful. I think that's the start of the process that will get us going to let us know what we have to do to provide all of our youngsters with an appropriate education.

Chairman KILDEE. Mr. Goodling?

Mr. GOODLING. Only to say, Mr. Deputy Secretary, that I'm glad I never used your name in my town meetings. Because in my town meetings I get all of these people who say, "Well, Peter Gray said so-and-so" and "Lee Iaccoca said so-and-so."

I've always said, "Well, why don't they give up their big jobs and come down here and serve." Then my Governor offered one a chance to come without even having to stand election, and he refused it. I'm glad I never used your name, and I thank you for coming to serve. We need you.

Mr. KEARNS. Bill, you're very nice to say that. I had an opportunity to make a lot of money over a long period of time, and I'm delighted to have an opportunity to serve the government.

Mr. GOODLING. Thank you.

Chairman KILDEE. Mr. Reed?

Mr. REED. Thanks, Mr. Chairman.

Mr. Secretary, just for my information and perhaps the record: internationally, do our major economic competitors have well thought out national standards for their national programs?

Mr. KEARNS. The answer to that is it varies substantially by country.

But on standards, Franci, you may know more about this than I do by country, and give me some help with this.

But most of the countries, first of all, have national systems, which we don't have, which is a lot of what we have been talking about here today. So it is easier for them to do that, but they do set out standards in different subject areas.

Franci?

Ms. ALEXANDER. Yes, they do substantially, every single one. The most recent in Europe was the addition of Great Britain as sort of the last hold out on the continent, and they have a national standard.

Mr. REED. I spent about 3 years in the Airborne, and I can't hear anything anymore, so could you translate? All those C-130s.

Mr. KEARNS. I was not in the Airborne, and I don't hear very well either, Mr. Reed. Franci said that basically all countries do, and the latest to add to this was the United Kingdom.

Mr. REED. Mr. Secretary, it seems that in a discussion of testing there are a couple of premises. One premise I would like you to comment upon is that there will be an instrumental effect; that simply adopting a system of testing will change the behaviors of children. Again, is that a premise that you're putting a lot of weight on or——

Mr. KEARNS. I don't put any weight on it whatsoever. I think it goes back to what we were talking about before. The reason to assess something and to understand where it is, is to know whether we are proceeding towards a goal. If you are not proceeding, you have to understand—it's like the total quality process. You have to understand the process that gets you to the end result.

You measure customer satisfaction from your customers and how they feel about your products and your service so that you can go back and understand the entire process as it goes through so you are fulfilling those needs. That's the reason for the assessment.

It goes back to the conversation again that the chairman and I and Mrs. Lowey had earlier, when we were talking about, you use the assessment to understand whether the system is providing for our youngsters the appropriate education. If it is not, then you can take actions. If there is an assessment system then in place, you will know whether those things are working or not.

Mr. REED. Again, I think it's important before we go forward to understand the premises. Your basic position is that you really don't care about the instrumental effect of testing as far as your proposals go, that it may or may not have an effect, but that's not what is driving the system?

Mr. KEARNS. What drives the system is to understand whether or not community by community and school by school we are providing a world-class education system for our children. If we are not, then we should fix it.

Mr. REED. Let me ask you another question, then. We have to make some very critical strategic decisions about deploying scarce resources. If we go down this road of standard-setting and then the follow-up, which is testing of the students, I think we should ask ourselves right now, is that the best strategic approach?

There are alternatives. For example, do we try to provide training assistance to teachers, and try to focus on teachers in terms of supporting their efforts to be better teachers in either a regional or State-by-State way?

I think there is a question of what is the best approach to reach the ultimate objective, which is an improved, enhanced educational system. Are you thinking about those strategic choices; and, if so, why are you keeping it to yourself, as it seems, to a testing regime of students?

Mr. KEARNS. Well, I don't think we're keeping it to ourselves. First of all, I think there are a lot of things that we need to do, a lot of transforming of ideas. I believe that the critical one, and I said this at the beginning, the critical ingredient is high standards

and assessments so you know where you are going, what kind of teacher training do we need.

There is no question we need teacher training and there will be dollars that will be required to particularly develop the particular testing mechanisms that are appropriate as we get into performance assessments and other types of things. We spend a lot of money on testing now. There will be up-front dollars required.

Over the longer period of time, we do a lot of testing. As you're going to hear from experts about this later, in this Nation, we probably overtest. Over a period of time, these should, the new assessments should, replace old assessments; but there will be some up-front dollars.

Out of the total number of dollars that we spend in education, $240 billion in elementary and secondary schools on a yearly basis, this is a relatively small amount of money. I think it is important.

By the way, we should not wait to start spending money on appropriate kinds of teacher training, and try to do that. That's one of the things that I spoke about before, the work that we're doing with NSF, particularly in math and science where the Federal Government does have a role, is to work on a strategy that improves teacher training not just the new teachers but in-service training.

There are 1.3 million K through 6 teachers out there. Most of them do not have the skills to bring our youngsters up to the level that they need to be at in math and science, and they need our support in training.

Mr. REED. I would ask a final question. You ran a very large organization and you probably had a very elaborate, I would assume, assessment program. Is that assessment program something that you see would be applicable to the public education in the United States, i.e., performance testing versus standardized testing versus training instructors, or evaluating the instructors and trainers? I mean, how does your experience in that——

Mr. KEARNS. I have very strong feelings, but I won't repeat what I just said to you before. High expectation levels are a basic driver. You have to understand those. Then you need to put the mechanisms in to assess that, and then basically you change the whole process of running the business.

I think beyond that concept. I think there are probably some things to learn from business, but I don't think that business itself has a lot to bring in the development of new testing mechanisms and new assessment processes, and so forth and so on.

Mr. REED. Did you test your employees periodically at Xerox?

Mr. KEARNS. Yes.

Mr. REED. Why did you do that? I mean, to see where you were going or to make them be better employees?

Mr. KEARNS. Well, when we went through the—when we put in place the total quality management process in 1982 and 1983, when we were on our way out of business, set new high expectation levels for the end of the decade, one of the things that we all agreed on—and we didn't have much money and we had to carve it out of others—was to make $175 million investment in training.

We did do some testing to find out along the way whether or not that training was effective. We had to go back and get into ele-

ments of statistics, for example, and we did some testing along the line to see whether people, in fact, were learning appropriately.

Mr. REED. But that was essentially to validate the training not as a constant sort of standardized approach to your workforce?

Mr. KEARNS. Well, we developed a lot of things. We did something that the top 25 of us had a lot of trouble with initially. One of the things in the process, we described new types of management behavior. We pounded it out, the 25 of us, to agree that we would let our subordinates actually measure our behavioral change, whether in fact we were doing that or not.

We described what was a role model, what was competent, and what needed work. Then we continually changed that as we went along and as we got better, but we agreed to let our subordinates measure that. It was a tool and a vehicle to see whether, in fact, we were doing it. It was also a motivator.

Mr. REED. Thank you, Mr. Secretary.

Thank you, Mr. Chairman.

Chairman KILDEE. Mr. Sawyer?

Mr. SAWYER. Thank you, Mr. Chairman.

Mr. Secretary, it has been reassuring to watch you this morning circle in on some topics that really don't want to sit still very easily. I don't mean to suggest by that that you are in any way dodging, because some of these things don't have one answer.

I want to talk about an element in all of this that I think is important, and it is the way in which we go about getting from here to there; agreeing as we do that getting from here to there is important. One of the very real problems that we face is to assure that we are measuring what we *think* we are measuring.

In asking this question, I want to stipulate a couple of things, if it's okay. First of all, I agree that there are some things we have to do now—we can't wait until other problems are solved to get ahead with something else.

Second, I would agree particularly with the comment that you made in your testimony about the danger of narrowing curriculum. We do not want to do that.

Third, I would agree about the importance of recognizing the diversity of the Nation as we attempt to go about all of this.

Having said that, I think I heard you talking about the importance of starting with standards, moving toward the development of assessment instruments and then toward remediation. It seems to me that the model that the National Council of the Teachers of Mathematics have provided is a more difficult, but far more sustainable, pathway to get from here to where we want to be.

That is, that they began with a set of specific objective contents designed to be built into the development of curriculum and the materials and the techniques necessary to teach those objective content standards. I believe what NCTM envisioned was that the actual standards for achievement they wanted would flow from that.

Having done that, they could begin to look for assessment instruments that they would validate for fairness and reliability and all of the things that you've spoken of and with that come both to the outcomes that they are looking for, both for expectations of accountability of entire systems and the capacity to remediate for

students and specific classrooms. That is a clearly much more sequential rather than simultaneous pathway of getting from here to there.

It is enormously important, I think, to come to decisions about what pathway we follow if we think that it is important to understand that we're measuring what we think we're measuring.

Having said all that, can you talk to us for a moment about the sequential versus simultaneous character of the development of standards and instruments and how important it is to follow the National Council of Teachers of Mathematics in the way in which they have gone about doing more than one thing, in getting from here to there?

Mr. KEARNS. Mr. Sawyer, let me comment. I'm going to ask Franci to come up here and help me with this.

Ms. ALEXANDER. Sure.

Mr. SAWYER. I really mean it to be a more simple question than it sounded in its phrasing.

Mr. KEARNS. Well, let me——

Mr. SAWYER. It is a question about simultaneity versus sequential approaches to getting from here to there.

Mr. KEARNS. Let me say that we can work on the standards. I'm going to let Franci talk a little bit about process with the National Council of Teachers of Mathematics.

But let me first—we chatted about this a little bit earlier. We can work on the standards and that work is going on in three areas now: in math, sciences, history, and, hopefully soon, geography, and then we're trying to work with different English groups, and then we're starting some support in the arts.

At the same time there is a lot of good work already taking place out there about new methods of assessment. In other words, that work is in fact taking place and there are people that are sitting behind me that are doing some of that work right now; they are out doing that. So you don't have to wait. In other words, that work, in my opinion, can take place.

What I think you do have to wait on, delayed on, as you start to use the new methods of assessment, I think we do have to be cautious. At some point we want to get high stakes or consequences, in my opinion, into it. My friend Al Shanker talks about that, is that at some point there ought to be some consequences. If you don't graduate from high school with an appropriate degree, you should not be able to go to college.

A system has to be developed as you get along, and we talked about earlier, starting in some of the lower grades with almost what I would call "no consequences tests" for the individuals, and we ought to be talking about, as Tom Petri or someone said earlier, what does that mean for the strategy in the school rather than for the student. I believe there is a lot that can take place——

Mr. SAWYER. I agree.

Mr. KEARNS. [continuing] in parallel without waiting. I do think we have to be cautious and there was a lot of discussion about this on the council, about so-called high stakes tests too early before we really understood were they valid, were they fair, and were we able to make appropriate comparisons.

Now, Franci, you might help me a bit?

Ms. ALEXANDER. Yes, and I think on the point of, should it be serial or simultaneous, we are in a situation where it is already simultaneous. Because, as David Kearns indicated, many States have already responded to the NCTM standards and have developed new assessments that reflect those standards. They didn't wait, so to speak, for the National Council of Teachers to continue developing standards or performance standards or work it through; they just went ahead and did it.

And for this reason—and I just came from a State 7 months ago—that's because the teachers said the biggest impediment to our implementing these new national standards for mathematics are the current tests so they said, "I'm not going to change my teaching if you're going to measure us and hold us accountable on the old tests." They said, "Give us a new way of showing it, and we'll do it."

After meeting with hundreds, thousands of teachers, we realized you have to do everything at once, so to speak. The very same thing happened with writing. That's a good example because we have a national writing project in this country and a lot of agreement about how writing should be taught. However, we found that many of the teachers had not changed classroom instruction, and that was because the test they were giving was an editing test.

When the States said, "Now we're going to have direct writing assessment, and kids will do an essay," we have a study of the teachers that showed immediately teachers started assigning real writing, changing their way of teaching, even though they had known for years before that what the new standards were. That's why it sort of needs to be everything at once.

Mr. KEARNS. Mr. Chairman, this is Franci Alexander, she is the—I know you know her, Dale, but is the deputy assistant secretary in OERI and was the executive director for the National Council on Education Standards and Testing.

Excuse me, sir.

Mr. KEARNS. Could you go back to the final point that you were making? I apologize.

Ms. ALEXANDER. Well, the point being that things have to happen somewhat at the same time because, indeed, for classroom behaviors—and that's what we all care about, what's happening in the classrooms—to change, you have to have a sense of what the new direction is, and teachers have to have confidence that they are going to be measured against those new standards.

Mr. SAWYER. If I hear what you're saying, then, to take a complex question and a complex answer, to reduce it to the simpler question that I was asking, if you think that the key to making simultaneity work is having gradually, increasing stakes rather than to begin with high stakes consequences?

Ms. ALEXANDER. Absolutely. In fact, the first test given in California had no consequences for individual students. For the first time, it will be another year before an individual student consequence is attached. And that is now after several years of writing the essays and using the new testing.

Mr. SAWYER. If we were to plot out a series of activities between now and some point in the future, one of the changing elements in

that would be the elevation of stakes as opposed to trying to make sequential the series of steps that we would go through?

Ms. ALEXANDER. That's correct.

Mr. SAWYER. That is absolutely the most coherent answer that I've gotten to that question that I've asked a number of times in a number of different ways, and I really thank you both for that. It was great.

Chairman KILDEE. Thank you very much.

David, this has been a very, very good hearing. I think we have been advanced a great deal on this. I think our discussions have been good. We are trying to put together a Title II, probably to the H.R. 4323, which will encompass many of the ideas that we've discussed here this morning. I really think that we have all gained from your testimony here this morning. I am grateful for that.

If you have any closing statement——

Mr. KEARNS. I am about wrung out.

Chairman KILDEE. Okay.

Mr. KEARNS. Franci, capped it off for Mr. Sawyer, so I'm going to leave it at that.

Chairman KILDEE. Very good. Well, thank you very much. I thank both of you for your appearance here.

Mr. KEARNS. Thank you, sir.

Chairman KILDEE. Our next panel will consist of Ms. Norma V. Cantu, director of Education Programs, Mexican American Legal Defense and Educational Fund, San Antonio, Texas; Dr. Leonard Rezmierski, superintendent of Northville Schools, Northville, Michigan, the Council for Exceptional Children; and Dr. H. D. Hoover, director of Iowa Basic Skills Testing, Iowa City, Iowa.

We welcome you here this morning, and you probably can proceed in the order in which I called your names.

STATEMENTS OF NORMA V. CANTU, DIRECTOR FOR EDUCATION PROGRAMS, MEXICAN AMERICAN LEGAL DEFENSE AND EDUCATIONAL FUND, SAN ANTONIO, TX; LEONARD REZMIERSKI, SUPERINTENDENT OF NORTHVILLE SCHOOLS, NORTHVILLE, MI, THE COUNCIL FOR EXCEPTIONAL CHILDREN; AND H. D. HOOVER, DIRECTOR, IOWA BASIC SKILLS TESTING, IOWA CITY, IA

Ms. CANTU. Thank you, Mr. Chairman.

I want to summarize my testimony, but first let me thank the subcommittee for the opportunity to be here.

If I were to tell you that I am the result of high expectations and high standards because my mother wanted me to succeed as a student, I would be lying. I mean, it's true that the parents set expectations of their children, but I came out of a very low-income part of the country, one of the three poorest SMAs in the United States, finished high school at the age of 16, was still a teen-ager when I enrolled at Harvard Law School, have been practicing law for 14 years.

It is not due to high expectations, it is due to exceptional opportunities that have been provided to me. Parents who enrolled me in prekindergarten at the age of 5, before there was a Head Start. An 8th grade English teacher who tutored me after school, followed

me all the way though the 12th grade, and gave me 2 hours of her unpaid time to work with someone that she thought had potential. A public school teacher did that for me. Parents who took me to the library twice a week so that I could get every book I could possibly check out.

Those are factors that contributed to my ability to be here today to talk to you in these few minutes. It wasn't high expectations alone, it wasn't high standards alone. That is the summary of the testimony I want to give you, that setting high expectations and expecting the job to be done would do a tremendous disservice to this country because it is not the problem of high expectations.

I think as a country we believe we are the best in the world. I think we have three problems: one is resources, the other is trained personnel, and the last is attitudes. Resources are a problem because in this country we have school finance problems.

For example, in the Los Angeles Public Schools, low-income children in east L.A. are getting $400 less per child spent on them. This translates to more than $12,000 less per classroom than our more affluent children in west L.A. School children in Texas get $1,000 per child less spent on them. This translates to $25,000 per classroom that low-income children have less access to. We have overcrowded classrooms, we have underpaid teachers, we have resource problems.

In terms of trained personnel, this affects the limited-English population the most because they need teachers who can make their curriculum understandable and comprehensible to them.

In many States, our systems allow waivers. They set the high standards and expect the teachers to be able to address the needs of the children and then give waivers away. We have teachers in California who are on their seventh year of waivers. The high standards are there, but the States then excuse or give the teachers a pass: the standards are waived.

There are no waivers for the limited-English-proficient students. They are still expected to perform and they are still expected to meet all of the requirements in order to graduate, but there are waivers to the school districts and the local administrations who are not brought—who are not made accountable for the fact that the children aren't learning and they should be.

The last problem is of attitudes. In my job at MALDEF, I run into administrators all the time. I run into a school principal who asks me, "If you think that Spanish-speaking child is so gifted and belongs in a gifted program, why isn't that child speaking English? If you think the child can do math, why aren't you satisfied to wait until the child speaks English, and then we will teach math to that child."

I'm running into attitudes. We tried to pass a bill in Texas to raise the compulsory age for high school graduation with the idea being that the longer you hold the students in school the more opportunity to teach the students something to be able to be competitive in the workplace. All but one school superintendent opposed raising the compulsory age. They wanted to get rid of students in the public schools as soon as possible. "Let's get them out." I mean, this is contraindicative to raising standards.

The children with disabilities who are also Spanish-speaking are being dumped in ways that—the assessments tools are there. They can be identified as to whether they need language assistance or whether they need special education assistance or both. The assessments are there but because of attitudes, because people are just not understanding where the children should go, the assignment practices are abysmal.

I heard a question earlier about universities. University attitudes? There are many universities I definitely would not want running the public schools. They have a problem themselves admitting minority students.

There are public universities that are still white enclaves, and I would have a serious problem with turning over the operation of public schools to a number of these universities. Moreover, they don't reach certain parts of the State. In Texas, there is a lawsuit because the major university systems have virtually ignored the border and ethnic region of the State.

I would ask this committee to really look at the issue of standards and assessments from the point of view of, does it really reach the special populations? Does it really reach the limited-English-proficient population, for example?

The limited-English population is assessed at a much higher rate than the mainstream English-speaking population. Limited-English children are assessed even before they enter the public schools because they have older children who are in Head Start; so some children are assessed at ages 0 to 5.

In Texas, the limited-English children are tested, they are assessed in the 1st grade, in the 3rd grade, in the 5th grade, in the 7th grade, in the 9th, 10th, 11th, and 12th grade as seniors before they are allowed to graduate, as college sophomores, if they get that far.

We have a lot of assessment going on. If we're calling for even more assessment or if we're calling even for replacement of current assessment, we're not really doing anything to help those kids. They are being tested a lot, they are being assessed a lot.

I frequently give people the example of a patient in the hospital with the nurse coming in every 5 minutes to take the patient's temperature. Taking the patient's temperature every 5 minutes is not going to cure the patient's fever. All you're getting is a wonderful documentation of the fact that that patient has a fever. If what we're suggesting is to come in and take the temperature 200 times a day, that patient will still have a fever.

There are many successful programs, they do need Federal support. We heard this morning from some examples from some universities that I would allow to operate and to participate in the public schools. The statistics are pretty unreal. I'm glad that you all are looking at them with a little bit of cynicism, and I would too.

I will close by giving you one last example of assessment. I'm a member of the Texas Bar and the California Bar. When I took the California Bar, they required a Professional Responsibility Exam, a standardized test with multiple choice to demonstrate that I had the potential to be an ethical attorney. I filled in all my bubbles

and scored in the nineties. I called my mother and I told her, "I'm an ethical attorney."

My mother told me two things. First, "Norma, it doesn't count as if you had gone to Mass on Sunday." Secondly, she said, "Norma, look who you're comparing yourself to."

[Laughter.]

Ms. CANTU. My mother who has always set high expectations of me, I think would require high expectations of this committee as well. If you have any questions either at the close of all of our presentations or at this time, I would be happy to answer them.

[The prepared statement of Norma V. Cantu follows:]

Mexican American
Legal Defense
and Educational Fund

733 15th Street, N.W.
Suite 920
Washington, D.C. 20005
(202) 628-4074
FAX (202) 393-4206

MALDEF

TESTIMONY OF THE
MEXICAN AMERICAN LEGAL DEFENSE
AND EDUCATIONAL FUND, INC. (MALDEF)

ON THE REPORT OF THE NATIONAL COUNCIL
ON EDUCATION STANDARDS AND TESTING (NCEST)

BEFORE THE SUBCOMMITTEE ON ELEMENTARY, SECONDARY AND
VOCATIONAL EDUCATION OF THE
COMMITTEE ON EDUCATION AND LABOR
U.S. HOUSE OF REPRESENTATIVES

PRESENTED BY NORMA V. CANTU
NATIONAL DIRECTOR, EDUCATION PROGRAMS
MALDEF

MARCH 18, 1992

National Office
634 South Spring Street
11th Floor
Los Angeles, CA 90014
(213) 629-2512
FAX (213) 629-0266

Regional Offices
542 South Dearborn Street
Suite 750
Chicago, IL 60605
(312) 427-8063
FAX (312) 427-8393

182 Second Street
2nd Floor
San Francisco, CA 94105
(415) 543-5598
FAX (415) 543-6235

The Block Building
140 E. Houston Street
Suite 300
San Antonio, TX 78205
(512) 224-5476
FAX (512) 224-5382

733 15th Street, N.W.
Suite 920
Washington, D.C. 20005
(202) 628-4074
FAX (202) 393-4206

Contributions Are Tax Deductible

52-840 O - 92 - 8

Testimony of MALDEF before the Subcommittee on Elementary, Secondary and Vocational Education

U.S. House of Representatives

March 18, 1992
Washington, D.C.

On behalf of MALDEF, I appreciate the opportunity to testify on the consequences on limited English proficient children of the report of the National Council on Education Standards and Testing (NCEST).

Background

The limited English proficient student population in this country is a rapidly growing group. In California, for example, close to one in four public school students is language minority. 1 In California, the number of limited English proficient students has tripled in one decade. In 1989, there were more than 740,000 limited English proficient students in California. Roughly two-thirds of those students speak Spanish as their dominant language, while the others speak Vietnamese, Cantonese, Cambodian, Filipino/Tagalog, Hmong, Korean, Lao, Mandarin, Japanese, and others. 2

MALDEF has litigated on behalf of the educational rights of the limited English proficient students in Texas, California, Idaho, Colorado and Illinois. In each instance, the courts have affirmed the U.S. Supreme Court in Lau v. Nichols that identical educational treatment of non-identical children could prove a violation of the children's rights. 3 In plain words, because LEP children are different from language majority children, they have a legal right to a different form of education, one that is relevant and compatible to their educational needs.

The scope of that legal right to compatible educational approaches was given more detail in the Fifth Circuit decision in Castaneda v. Pickard. 4 The holding of the Castaneda case involved a four-prong test: a school district must (a) develop a pedagogically sound approach to address the English language and access needs of its LEP children, (b) implement that approach by allocating necessary resources to assure its effectiveness, (c) develop an evaluation system to assure that it is working, and (d) respond affirmatively to the evaluation when it shows that the students are not meeting with success.

MALDEF and other advocacy groups are still litigating to require districts to comply with prongs (a) and (b) of the Castaneda standard. We are aware of many districts with problems in (c) not developing assessment systems and (d) not responding to their failures with regards to LEP students.

The Good News

Given that background, how does the NCEST report affect language minority students? There are several foreseeable consequences of the report, some negative and some positive.

The good news is that the report affirms what educators have been preaching for a number of years, i.e. that students respond well to high expectations. Thus, the raising of standards, when done in a positive manner that bolsters students' self-esteem, can be productive for reducing the numbers of students who drop out of school. 5

Secondly, the other good news is that the report refrains from adding another layer of testing to the already-overburdened school populations who suffer from too many tests. 6

Thirdly, the NCEST report repudiates having a single national achievement test. This is consistent with the outcome of study performed by the National Commission on Testing and Public Policy, strongly recommending that test scores are imperfect measures and should not be used alone to make important decisions about the allocation of educational opportunities. 7

Finally, the NCEST report suggests that the role of the assessment information should be a supportive one, not to label or brand students, but to provide continuous feedback on the status of our educational system.

The Bad News

The bad news is that despite the assurances from the makers of the NCEST report that no national test in intended, the local educators will, in fact, administer tests reflecting the national standard. In this country, limited English proficient students have historically encountered racism and neglect in the area of assessments, so there is much room for skepticism of any national proposal that focuses on national standards and assessments. Historically, there has been an uncritical acceptance of tests for the purposes of prediction, tracking, grouping and other forms of differentiation that arbitrarily selected children for a lower slot in a hierarch of educational opportunities and privileges. 8

Because the limited English proficient population is concentrated in certain geographic areas of the United States and because the population has certain needs, the population is easy to ignore by majority policymakers. Thus, limited English proficient students will very likely be negatively affected by the national standards.

MALDEF is not convinced that any national standard, set in Washington, is going to change the attitudes of the people who actually write the curriculum and evaluate the local schools.

Please let me share with you one example from my personal experience as an advocate for LEP students. In 1985, in the San Rafael public schools in California, one of the third-grade students was being pulled out of his class every day to report to the first grade classroom to assist the first grade teacher. Because the third-grade student understood both English and Spanish, he was being used as a translator for the first grade teacher who was monolingual English speaking. The principal was aware of the fact that the third grader was missing his classes. He knew that the third grader was doing the actual teaching in the first grade class. In fact, the principal thought so highly of the third grader that he would call him into the office to ask for his help in translating so that he could "card," ask for immigration papers, of new students enrolling in his school. The third grader finally objected to the role that he was being required to play. He complained to his mother and through community pressure and the threat of a lawsuit, San Rafael changed its policies and hired its first bilingual teacher. The third grader was returned to his regular class, where he could resume working as a third grader, not an unpaid teacher's aide.

The weakness of the NCEST report as it relates to the LEP population is that it assumes that each of the states is committed to accurate assessments of LEP students. They are not.

The NCEST report acknowledges that there is a need for equitable treatment of students, but makes no solid recommendations for addressing the lack of Hispanic role models in the schools, the lack of appropriate language program, the high pupil-teacher ratios and the limited resources available to the schools in which Hispanic students enroll.

The NCEST report assumes that there will be changes in educational curriculum and instruction as a result of the heightened standards. This is not the case. In Texas, for example, the cut-off scores for the high school graduation exams have been raised every year for the past four years. The heightened standards have not raised student performance. To the contrary, the heightened standards have produced more student failures in the form of students who will be denied the regular high school diploma. Instead, students will be receiving a certificate of attendance, a stigma that will follow them for their entire lifetime.

Conclusion

Researchers have noted numerous differences in how limited English proficient students learn, as compared to language majority children. For example, limited English proficient students learn verbally, not through reading. 9 The limited English proficient students exhibit cooperation values, not competitive values. 10 Thus, it be would unwise to assume that limited English proficient students would be spurred to compete on an international basis,

when they have not have experiences in competing at local levels. Finally, limited English proficient students share many of the characteristics of at-risk students.

In conclusion, the producers of the report wrongly conclude that the American public is satisfied with the current performance of students. We are clearly not. But the dissatisfaction with student performance cannot be cured unless there is a commitment of resources. With specific regard to LEP students, the frustration that many educators have voiced is that there is a compulsive focusing on refining language assessment instruments. 11 MALDEF recommends that the focus should be on increasing the numbers of students served by language programs and the amount of time that students spend in those programs.

You have no doubt heard the figures on the dependency ratios, but it is useful to hear them again. The dependency ratio for children under 18 now stands at 42 children per 100 workers. The dependency rates for persons over the age of 65 is now 19 per 100 workers. 12 By 2030, there will be an equal number of children and retirees. By 2030, the dependency ratio will be 74 per 100 workers. The burden in the next century will fall on Hispanic workers and other minority groups. 13

For this reason, we urge the Subcommittee to continue to study the report in view of the factors mentioned in this testimony.

S O U R C E S

1. Olsen, L. and Mullen, N., Embracing Diversity: A Publication of the California Tomorrow Immigrant Students Project, California Tomorrow (1990).

2. Id.

3. Lau v. Nichols, 414 U.S. 563 (1974).

4. Castaneda v. Pickard, 648 F.2d 989 (5th Cir. 1981); Gomez v. Illinois State Board of Education, 811 F.2d 1030 (7th Cir. 1987); Keyes v. School District No. 1, Denver, Colorado, 576 F.Supp. 1503 (D.Colo. 1983).

5. Orum, L.S., Hispanic Dropouts: Community Responses, Office of Research, Advocacy and Legislation, National Council of La Raza, Washington, D.C., (July 1984).

6. From Gatekeeper to Gateway: Transforming Testing in America, National Commission on Testing and Public Policy (1990).

7. Id.

8. Gonzalez, Gilbert G., Chicano Education in the Era of Segregation, Associated University Presses (1990), at 76.

9. Wells, S.E., At Risk Youth: Identification, Programs and Recommendations, Teacher Ideas Press (1990).

10. Bejamo, Y., "A Cooperative Small Group methodology in the Language Classroom," TESOL QUARTERLY 21, 483-504 (1987); Jacob, E. and Mattson, B., "Cooperative Learing: Instructing LEP Students in Heterogeneous Classes."

10. Padilla, A.M. (ed.), Bilingual Education: Issues and Strategies, Sage Publishers (1990).

11. "Here They Come: Ready or Not. Who's Dependent on Whom?," Educatıcn Week, May 14, 1986), p. 14.

12. Hayes-Bautista, D.E., Schink, W.D, and Chapa, J., The Burden of Support: Young Latinos in an Aging Society, Stanford University Press, Stanford, CA (1988).

13. Id.

Chairman KILDEE. Thank you very much for your testimony. I think we will go to Dr. Rezmierski first and then Mr. Hoover, and then we will ask questions. Thank you.

Ms. CANTU. Thank you.

Mr. REZMIERSKI. [Speaking Polish]

Chairman KILDEE. [Speaking Polish]

Mr. REZMIERSKI. [Speaking Polish]

Mr. Chairman, I am Leonard Rezmierski, and I am currently superintendent for the Northville Public Schools in Northville, Michigan. I also currently hold the position of Governor at Large, U.S. in the leadership of The Council for Exceptional Children. I will paraphrase from our testimony aspects of this document.

We thank the Chairman and the distinguished members of this House subcommittee for the opportunity to testify regarding the past and future work of the National Council on Education Standards and Testing. The Council for Exceptional Children, representing over 52,000 professionals and others concerned with education of infants, toddlers, children and youth with disabilities as well as children and youth who are gifted, strongly supports improving educational outcomes for all of America's students.

Our membership is comprised of professionals from many disciplines who provide a unique knowledge base from which to offer comment and recommendations. For the past several years, beginning with the 1983 report, "A Nation At Risk," there has been a nationwide focus on restructuring and improving our Nation's schools. More recently, the call for world-class standards has gathered momentum. When President Bush announced his reform plans in America 2000, discussion about national standards and national tests intensified.

We commend the Congress for passing the legislation which brought together the members of the National Council on Educational Standards and Testing. The deliberations and report of NCEST bring forth many of the sensitive issues involved in establishing national standards in testing.

CEC has followed with interest in deliberations of NCEST. We believe that they thoughtfully address the questions asked of them in the authorizing legislation as well as in the September 23, 1991, letter from you, Mr. Chairman, in conjunction with your colleagues, Representative Goodling and Senator Hatch.

We were especially pleased with your September 23 letter, since it addressed many specific concerns which we also share, such as how national standards and tests would impact on educationally disadvantaged children, children with disabilities, and children with limited English proficiency.

We agree with the Congress that if we are to have national standards and national tests, they must be inclusive of all America's students. There should be no misunderstanding that all students include students with disabilities and other students who may experience learning difficulties.

We will focus, or I will, the rest of my testimony how the three key areas of the NCEST report would impact students with exceptionalities: first, world-class standards; second, national assessments; and, third, a coordinating structure.

World-class standards: Setting the world-class standards causes us more than a little anxiety since we believe that if national standards are to be set, they must be fair to all students. We strongly believe that educational outcomes for all students need to be improved. Our very strong concern, however, is that the setting of standards is arbitrary. There is no empirical evidence that can tell us what world-class standards are and what students should know and when.

Furthermore, given that various professional organizations are involved in setting standards, there is no policy which requires these groups to set standards which will challenge our most able students as well as those with special learning needs. The standards must be rigorous enough for our students who are gifted, challenging enough for the majority of our students, and flexible yet challenging for students with disabilities.

The example I would cite from our own district would involve gifted students, who will go beyond the proficiency in the foreign languages such as Spanish to learning Mandarin Chinese, to our most handicapped children who learn up to 100 basal words in their language board or their WOLF System. We must include all the students.

The issue of including all students in performance standards yet ensuring that performance standards are realistic for all students must be further studied. The Standards Task Force of NCEST propose one possible approach by having a scale of student performance standards.

Another alternative may be to view standards in the core subjects as dimensional, and with the knowledge that development and learning occur incrementally over time; hence, students will achieve the standards at varying rates. Examples of this would be the concept of mastering learning-outcomes-based education.

As an alternative, we strongly encourage the acknowledgment that achievement of world-class standards in the core subjects by all students may not be appropriate nor relevant to their post-school activities.

As a result, we strongly encourage you to support the establishment of measurable standards which are relevant to everyone's post-school activities; namely, standards that focus on ensuring that all students learn to use their minds well so they will be prepared for responsible citizenship, further learning, and productive employment as stated in the National Education Goal 3.

CEC was pleased that NCEST recommended that standards be set in areas other than the five core subjects. We strongly agree with this recommendation. We are concerned that if standards are set only in English, math, history, science, and geography, such a baseline may lead to a narrowing of the curriculum which would have quite a negative impact on students with disabilities as well as those who are gifted.

When working with students with severe disabilities, it may be more important to teach the student independent living or a vocational skill rather than algebra. Or other students, including students who are gifted, may also need a challenge of subjects outside the core subjects, such as foreign language, performing arts, com-

munication arts, and higher order thinking skills and problem-solving.

We encourage the subcommittee to stress the benefits of a diverse curriculum. If the members of the subcommittee support the development of national standards, we strongly agree with the NCEST recommendation that the standards be developed through a broad-based process that involves educators including teachers, parents, students—the ultimate client—business people, and the public.

If such a process is followed, those with expertise in working with students who are gifted as well as those with students who have disabilities can provide appropriate input. Our concern, however, is that such broad-based input will not be sought.

CEC strongly supports the recommendation of the NCEST that the school delivery standards and system performance standards be established. We cannot support the setting of standards for students without also requiring schools and school systems to be responsible for ensuring that all students have an opportunity to learn.

If students must meet world-class standards, school delivery standards must also be world-class. We must prefer the definition and accountability required of the school delivery standards as outlined in the Standards Task Force Report than the language adopted by NCEST in its final report.

Given the widely acknowledged inequality across America's public school systems, we contend that it would be unfair and impossible to hold students to the same set of performance and content standards without seeking to ensure that equity is present through uniform delivery standards. All students must have an equal opportunity to learn and to improve performance.

The second area of assessment: CEC has grave misgivings about the development of national assessments. However, if national assessments are to be developed, the system must include, again, all students. Much of our concern stems from the fact that many States currently exclude many students with disabilities from testing.

Part of the reason for excluding students with disabilities is that States have technical problems assessing this population, but another part of the reason for excluding such students is the desire to raise the State's overall test scores. Given that an estimated 4.5 million students receive special education services in our country, the practice of exclusion cannot be allowed in any type of assessment system.

The National Center on Educational Outcomes at the University of Minnesota, funded by the U.S. Department of Education, is working to define domains of learning and develop a model of desired educational outcomes and a system of outcome indicators that apply not only to students with disabilities but to all American students.

This model will provide a progressive approach for looking at students' performance in a truly inclusive way of assessing the process of all students. One of the ongoing activities of the National Center is to identify technical issues that need to be addressed

as policymakers move toward greater assessment of educational outcomes for students with disabilities.

Let me quote. Some of their findings are discussed in the "Report on the Inclusion of Students with Disabilities in National and State Data Collection Systems," and are quite disturbing. Let me cite just a couple.

As currently designed, most of the existing State and national data collection programs exclude large portions of student population with disabilities. At the national level it is estimated that approximately 40 to 50 percent of all school-age students with disabilities are excluded from the most prominent and national data collection programs. Another finding is that reasons typically given for exclusion of students with disabilities range from concerns about providing proper accommodations to concerns about the potential aversiveness of the assessment situation for the student.

Finally from the report, the exclusion of students with disabilities results in significant problems in obtaining representative samples. This, in turn, creates difficulties in estimating national and State level statistics and in completing accurate policy studies. The findings of the National Center are important to consider if we are to develop a national assessment system.

Perhaps the most valuable long-term contribution that this national center can offer us is the work which it is doing in the area of identifying educational outcomes for all students. The center is working with approximately 200 stakeholder groups in an effort to develop consensus on a model of outcomes and a comprehensive system of indicators for students.

If we are to have a national assessment system, we strongly support the recommendation of NCEST that it be a system of multiple assessments linked to the national standards that will measure the progress of individuals, schools, school districts, States, and the Nation.

CEC agrees with the points expressed by the NCEST when addressing individual student assessments. However, we think it is important to add to this list that the individual assessments may need to be administered in more than one way and/or modified to meet the needs of those being assessed.

They must also address the broad range of skills required to succeed in work in community settings. Likewise, the assessments need to be open-ended enough to provide all students with opportunities to express the range of their knowledge and skills rather than being limited to what is being asked. The greatest problem we see right now is the fact that far too many of our students consider school time as prison time. The third and last area deals with the coordinating structure.

I would like to move for a second, having read all this information sent to me as a school superintendent and in sharing this information with others in my community, especially emphasis in this report on from page 35 on through 39, I find an interesting comment given to me by one community person after also reading it, saying we should market a bumper sticker stating, "Thank God we don't get all the government we pay for!" A bit of humor.

[Laughter.]

Mr. REZMIERSKI. CEC strongly urges the subcommittee to support the NCEST recommendation that a new entity, the National Education Standards and Assessments Council, be established to work with a reconfigured national educational goals panel to certify standards and criteria for assessment.

CEC was pleased with the membership and functions that NCEST recommended for such a new entity. Our concern with such an entity, however, is that at least one member of the Council have expertise in working with individuals with disabilities and those who are gifted.

We also wish to express a sense of urgency regarding the formation of the National Education Standards and Assessments Council, given that the Office of Education Research and Improvement has already awarded funding to some groups to develop world-class standards. Unless the new council is in place soon, it won't be able to begin advising and guiding the development of standards.

In closing, we wish to strongly emphasize that we believe the development of the standards and national assessments will not, in and of themselves, improve education in America. We urge the subcommittee to look closely at the pieces of systemic change which must be in place for the schools to meet the challenge of drastically improving the performance of all of our students especially those who have special learning needs.

In particular, teachers must be involved in the reform efforts both as contributors to what is needed in systemic change, i.e., curriculum development, but also as well as recipients of appropriate training both at the preservice and in-service levels.

We realize that these issues are very complex and will require continued study. We appreciate the time and attention you are taking to understand the implications of national standards and assessments and know that you will continue to provide strong congressional leadership as we work to improve educational outcomes for all students.

Thank you for this opportunity to share our concerns, and please know that we are available at all times to assist you in fulfillment of your legislative responsibilities.

Thank you, Mr. Chairman.

[The prepared statement of Leonard Rezmierski follows:]

STATEMENT OF DR. LEONARD REZMIERSKI, NORTHVILLE, MICHIGAN

Mr. Chairman, I am Leonard Rezmierski and I am currently superintendent for the Northville Public Schools in Northville, Michigan. I also currently hold the position of Governor at Large, U.S., in the leadership of The Council for Exceptional Children.

We thank the Chairman and the distinguished members of the House Subcommittee on Elementary, Secondary, and Vocational Education for the opportunity to testify regarding the past and future work of the National Council on Education Standards and Testing. The Council for Exceptional Children (CEC), representing over 52,000 professionals and others concerned with the education of infants, toddlers, children and youth with disabilities as well as children and youth who are gifted, strongly supports improving educational outcomes for all of America's students. Our membership is comprised of professionals from many disciplines who provide a unique knowledge base from which to offer comment and recommendations.

For the past several years, beginning with the 1983 report, *A Nation at Risk*, there has been a nationwide focus on restructuring and improving our Nation's schools. More recently, the call for "world class standards" has gathered momentum. When President Bush announced his reform plans in America 2000, the discussion about national standards and national tests intensified. We commend the Congress for passing the legislation which brought together the members of the National Council on Education Standards and Testing (NCEST). The deliberations and report of NCEST bring forth many of the sensitive issues involved in establishing national standards and testing.

CEC has followed with interest the deliberations of NCEST. We believe that they thoughtfully addressed the questions asked of them, both in the authorizing legislation as well as in the September 23, 1992, letter from you, Mr. Chairman, in conjunction with your colleagues, Representative Goodling and Senator Hatch. We were especially pleased with your September 23 letter since it addressed many specific concerns which we also share, such as how national standards and tests would impact on educationally disadvantaged children, children with disabilities, and children with limited English proficiency. We agree with the Congress that if we are to have national standards and national tests that they must be inclusive of all America's students. There should be no misunderstanding that "all students" includes students with disabilities and other students who may experience learning difficulties. Of particular concern to us are the special needs of students with disabilities, students who are gifted and students who are culturally diverse or whose primary language is not English.

We will focus the rest of our testimony on how three key areas of the NCEST report would impact on students with exceptionalities: world class standards, national assessments and a coordinating structure. We are pleased that there is an individual on the panel who will address the impact of these issues on students who are culturally and linguistically diverse.

World Class Standards

The setting of "world class standards" causes us more than a little anxiety since we believe that if national standards are to be set, they must be fair to all students. We strongly believe that educational outcomes for *all* students need to be improved. And, we strongly believe that if world class standards are to be set, they must be inclusive of all learners; therefore, challenging all students to improve their performance. As a result, we support the Council's intent in recommending the establishment of national standards as a means to "raise the ceiling for students who are currently above average and to lift the floor for those who now experience the least success in school, including those with special needs." Our very strong concern, however, is that the setting of standards is arbitrary—and there is no empirical evidence that can tell us what world class standards are and what students should know, and when. Furthermore, given that various professional organizations and groups are involved in the setting of standards, there is no policy which requires these groups to set standards which will challenge our most able students as well as those with special learning needs. There is much diversity among learning needs, styles and capabilities of students in our schools. As a result, we believe that the standards must be flexible in order to be realistic, in order for students to be appropriately challenged, and in order to ensure that improved learning is recognized. For instance, the standards must be rigorous enough for our students who are gifted, challenging enough for the majority of our students, and flexible yet challenging for students with disabilities. However, this does not mean that we support the differential treatment of students based on a student's ability to learn. The groups setting standards must address the issue of how world class standards will accommodate *all* students.

The issue of including all students in performance standards, yet ensuring that the performance standards are realistic for all students, must be further studied. The Standards Task Force of NCEST proposed one possible approach by having a "scale of student performance standards." Another alternative may be to view standards in the core subjects as dimensional and acknowledge that development and learning occur incrementally over time; hence, students will achieve the standards at varying rates. Further, if curriculum frameworks tied to the standards are in place, then teachers and students will have guidance in working towards the standards. We do not know if either of these proposals are the answer, but we do know, for example, that a gifted student can achieve a much higher content standard in mathematics than a student with a mental disability. However, we believe that both should be held to their individual "highest standard" to ensure that both students are challenged to learn. They both should be expected to achieve.

*Permit us to offer another example: Will students with severe disabilities be ex-*pected to meet world class standards in the five core subjects? If yes, will they be expected to reach the same standards as the majority of the school-age population? If not, what performance standards will they be expected to meet? We cannot accept an answer that students with severe disabilities are "exempt" from meeting world class standards.

As an alternative, we strongly encourage the acknowledgement that achievement of world class standards in the core subjects by all students may not be appropriate, nor relevant to their post-school activities. As a result, we strongly urge you to support the establishment of measurable standards that focus on ensuring "that all students learn to use their minds well so they will be prepared for responsible citizenship, further learning, and productive employment," as stated in National Education Goal Three.

We believe that one of the goals of the education system is to provide students with the "tools" needed to live productive, independent lives. The setting of measurable standards in these areas would support Goal Three, would be relevant to all students, and would help to make U.S. students competitive in our changing world. Examples of such standards are: standards for achievement of independent living, standards for vocational skills that lead to productive and secure employment, and standards for lifelong learning.

CEC was very pleased that NCEST recommended that standards be set in areas other than the five core subjects. We strongly agree with this recommendation. We are concerned that if standards are set only in English, mathematics, history, science, and geography, such a baseline may lead to a narrowing of the curriculum, which would have quite a negative impact on students with disabilities as well as those who are gifted. When working with students with severe disabilities, it may be more important to teach the student independent living skills or a vocational skill rather than algebra. Or, other students, including students who are gifted, may also need the challenge of subjects outside of the core subjects, such as foreign languages, performing arts, communication arts, and higher order thinking skills and problem-solving. We encourage the subcommittee to stress the benefits of a diverse curriculum.

If the members of the subcommittee support the development of national standards, we strongly agree with the NCEST recommendation that the standards be developed through a broad-based process that involves educators (including teachers), parents, students, business people and the public. If such a process is followed, those with expertise in working with students who are gifted as well as those who work with students with disabilities can provide input. Our concern, however, is that such broad-based input will not be sought. For example, it seems that everyone concerned with the setting of standards cites the work of the National Council of Teachers of Mathematics and the process that they followed. We have observed much deserved praise for the association and the work it has done, but we have not observed discussion respecting whether or not there should be a broad base of input if these are to become "national standards." Were individuals who teach students who are gifted or students with disabilities involved in the setting of these standards? If not, will there be an opportunity for "broad-based" input before the standards are declared "national"?

CEC strongly supports the recommendation of the NCEST that school delivery standards and system performance standards be established. We cannot support the setting of standards for students without also requiring schools and school systems to be responsible for ensuring that all students have an opportunity to learn.

CEC is disappointed that the NCEST allows for each State to select the criteria it will use for assessing a school's capacity and performance when determining the school delivery standards. We understand and support the NCEST desire to affirm the individual State responsibility for education, but we are concerned that such flexibility will allow States to have "weak" school delivery standards. If students must meet world class standards, school delivery standards must also be world class. We much prefer the definition and accountability required of the school delivery standards as outlined in the Standards Task Force report than the language adopted by the NCEST in its final report. Given the widely acknowledged inequality across America's public school systems, we contend that it would be unfair and impossible to hold students to the same set of performance and content standards without seeking to ensure that equity is present through uniform delivery standards. All students must have an equal opportunity to learn and to improve performance.

Assessment

CEC has grave misgivings about the development of national assessments. However, if national assessments are to be developed, the system must include *all* students. Much of our concern stems from the fact that many States currently exclude many students with disabilities from testing. Part of the reason for excluding students with disabilities is that States have technical problems assessing this population (i.e., providing alternative forms of assessment); but another part of the reason for excluding such students is the desire to raise the States' overall test scores. Given that an estimated 4.5 million students receive special education services the practice of exclusion cannot be allowed in any type of assessment system.

The National Center on Educational Outcomes (NCEO) at the University of Minnesota, funded by the U.S. Department of Education, is working to define domains of learning and develop a model of desired educational outcomes and a system of outcome indicators that apply not only to students with disabilities, but to all of America's students. This model will provide a progressive approach for looking at student performance and a truly inclusive way of assessing the progress of all students.

One of the ongoing activities of the National Center is to identify technical issues that need to be addressed as policymakers move toward greater assessment of educational outcomes for students with disabilities. While a variety of issues have been cited, the two most critical issues identified are inclusion/exclusion decisions and modification of tests. These two issues were also identified as critical facts when the Center looked at national and State data bases for existing information on students with disabilities. Some of the findings, discussed in the *Report on the Inclusion of Students with Disabilities in National and State Data Collection Systems*, are quite disturbing. Among the findings:

As currently designed, most of the existing State and national data collection programs exclude large portions of the student population with disabilities. At the national level, it is estimated that approximately 40 percent to 50 percent of all school-age students with disabilities are excluded from the most prominent national data collection programs (e.g., National Assessment of Educational Progress [NAEP], National Education Longitudinal Study [NELS]) that are playing a critical role in the evaluation of the current reform initiatives.

Reasons typically given for exclusion of students with disabilities range from concerns about providing proper accommodations to concerns about the potential aversiveness of the assessment situation for the student.

A sizeable portion of excluded students should not have been excluded from data collection programs and could readily participate (some with testing accommodations, others without), in such data collection programs.

The exclusion of students with disabilities from State and national data collection programs occurs at a number of different stages: (a) during the development of assessment instruments; (b) when the data are collected; and (c) during the analysis, interpretation, and reporting of the results.

The ability to extract useful national and State policy-relevant information on the outcomes of students with disabilities from national and State data collection programs is seriously hampered by the extensive exclusion of portions of this population. The exclusion of students with disabilities results in significant problems in obtaining representative samples. This, in turn, creates difficulties in estimating national- and State-level statistics (e.g., dropout rates) and in completing accurate policy studies.

The findings of the National Center are important to consider if we are to develop a national assessment system. First, if the purpose of the system is to help the Nation track our progress towards meeting the performance standards, all students must be included if we are to acquire an accurate and complete national portrait. Secondly, students with disabilities must be included in efforts to develop national assessments. Only if such students are included in test development will we learn what items or test administration procedures may need to be dropped or modified in order to accommodate any special needs.

Given the status and emphasis that is placed on the National Assessment of Educational Progress (NEAP), it is troubling that so many of our Nation's students are excluded from this test. If the NAEP can exclude 40 percent to 50 percent of students receiving special education support and yet be heralded as an effective indicator for measuring the Nation's educational progress, what guarantee do we have that a new national assessment system won't also be allowed to exclude large numbers of the student population?

Perhaps the most valuable long-term contribution that the National Center can offer us is the work which it is doing in the area of identifying educational outcomes for all students. The Center is working with approximately 200 stakeholder groups in an effort to develop consensus on a model of outcomes and a comprehensive system of indicators for students.

If we are to have a national assessment system, we strongly support the recommendation of the NCEST that it be a system of "multiple assessments linked to the national standards that will measure the progress of individuals, schools, districts, States, and the Nation." Furthermore, we support having two components: individual student assessments as well as large-scale *inclusive* sample assessments.

In addition, we believe that there must be formative as well as summative assessments. For example, the summative assessments proposed will tell us how many of our students are achieving the world class standards; formative assessments will tell us the extent to which students are making progress in relation to meeting the standards. Such formative assessments would provide us with data to be used in documenting improvement and in developing a strategy for working with students who have not yet achieved, or who have exceeded, the standards. Such assessments would have policy implications and would provide us with baseline information about how our students are currently performing in relation to the standards.

CEC agrees with the points expressed by the NCEST when addressing individual student assessments. However, we think it is important to add to this list that the individual assessments may need to be administered in more than one way and/or modified to meet the needs of those being assessed. They must also address the broad range of skills required to succeed in work and community settings.

There is a great need for, and we encourage the subcommittee to support, much more research in the area of alternative forms of assessment as well as in identifying strategies for modifying assessments to ensure that assessments can fairly and accurately determine what an individual knows. Likewise, the assessments need to be open-ended enough to provide all students with opportunities to express the range of their knowledge and skills rather than being limited by "what is asked."

Coordinating Structure

CEC strongly urges the subcommittee to support the NCEST recommendation that a new entity, the National Education Standards and Assessments Council, be established to work with a reconfigured National Education Goals Panel, to certify standards and criteria for assessments. CEC was pleased with the membership and functions that NCEST recommended for such a new entity. Our concern with such an entity, however, is that at least one member of the Council have expertise in working with individuals with disabilities and those who are gifted. We believe that this is an essential requirement if our schools are to achieve the accommodations mandated by the Congress through the Americans with Disabilities Act. Unless such expertise is represented on the Council, it would be possible for a set of standards or an assessment to be certified without anyone ever asking, "Is this inclusive of *all* students"?

We also wish to express a sense of urgency regarding the formation of the National Education Standards and Assessments Council. Given that the Office of Educational Research and Improvement has already awarded funding to some groups to develop world class standards, unless the new Council is in place soon, it won't be able to begin advising and guiding the development of the standards.

Closing Thoughts

In closing, we wish to strongly emphasize that we believe the development of standards and national assessments will not in and of themselves improve education in America. We urge the subcommittee to look closely at all of the pieces of systemic change which must be in place for the schools to meet the challenge of drastically improving the performance of all of our students, especially those who have special learning needs. In particular, teachers must be involved in the reform efforts, both as contributors to what is needed in systemic change (i.e., curriculum development, as decision-makers within the schools) as well as recipients of appropriate training, both at the preservice and inservice levels.

We strongly support the work of the Implementation Task Force as it tried to summarize what change needs to occur in the schools in order for all of America's students to have an equal opportunity to learn. We urge the subcommittee to review the work of the Implementation Task Force as you work to complete your education reform legislation. And, we would again emphasize the importance of achieving equity across the public schools to ensure that all students, especially those with special learning problems and needs (e.g., students who are economically disadvan-

taged, students with disabilities, and students who do not speak English as a first language), have an equal opportunity to improve their performance.

We realize that these issues are very complex and will require continued study. We appreciate the time and attention you are taking to understand the implications of national standards and assessments and know that you will continue to provide strong congressional leadership as we work to improve educational outcomes for all students. Thank you for the opportunity to share our concerns, and please know that we are available at all times to assist you in the fulfillment of your legislative responsibilities.

Chairman KILDEE. [Speaking Polish]

Mr. REZMIERSKI. [Speaking Polish]

Chairman KILDEE. Dr. Hoover?

Mr. HOOVER. [Speaking Polish]

As the director of a voluntary K-8 testing program, operated as a service to the schools of Iowa by the University of Iowa, I welcome the opportunity to address this committee regarding the appropriateness and feasibility of a national test. The University of Iowa has furnished tests and services to all schools in Iowa in grades 3 through 12 for over 50 years.

Currently, over 99 percent of the public and private schools in the State annually participate in the voluntary program I direct using the Iowa Tests of Basic Skills in grades K to 8, and a companion program for the high schools using the Iowa Tests of Educational Development.

Outside of the State of Iowa, the Iowa Tests are distributed by Riverside Publishing Company, a wholly-owned subsidiary of Houghton Mifflin. However, the University of Iowa holds the copyrights to these tests and the university-based authors not only operate the State programs in Iowa, but exercise editorial control over the content of the tests and associated materials and services. Neither the State department of education nor the State legislature has anything to do with either program.

We are convinced that the voluntary nature of these programs is why Iowa is the only State having extensive longitudinal information on the performance of students, as evidenced by the two CBO reports on achievement trends completed in the late 1980's, following the release of the report, "A Nation At Risk," in 1983.

The most recent update of these trends is appended to my testimony as Exhibit A. I have also appended as Exhibit B a table showing how the State of Iowa compares to the rest of the Nation in achievement based on data from Iowa Testing Programs, and independently, the National Assessment of Educational Progress, NAEP.

The purpose of these exhibits is to show that by using currently available tests we are capable of not only monitoring trends in achievement, but also of making the kinds of comparisons that led to the reporting of NAEP test results State by State.

One of the reasons I wanted to mention the Iowa Testing Programs was because of the amazing similarity of the original high school program started in 1929 to a number of the current proposals for a national test, including that of the National Council on Educational Standards and Testing, NCEST, in its recently released publication "Raising Standards for American Education."

The original Iowa high school program was an end-of-year, subject matter based, statewide academic contest employing a number

of types of tests; that is, they weren't all multiple choice. While this program is extremely popular with school administrators, policy-makers, and the press, it had many of the negative effects alluded to in the NCEST report regarding high stakes tests and was finally replaced in 1942 by a new program using the Iowa Tests of Educational Development, ITED, with the entire focus of the program on improving instruction for individual students.

The ITED emphasizes critical thinking skills and serves as a model for what a high school graduate is expected to know and be able to do. The ITED has also served as a model for a number of other well-known tests. In fact, the first forms of the Tests of General Educational Development, the GED, the National Merit Qualifying Test, and the American College Tests, the ACT, simply borrowed intact tests from the ITED battery to get started.

I would add here a side note. Anybody who believes you can't measure critical thinking skills with tests like this, I've got a copy of the ITED, and I would love to have the members of this committee take the tests. I would be more than willing to give them answer sheets, send them to Iowa City, and we will score those them.

Mr. SAWYER. Both of us?

Mr. HOOVER. It's a big crowd.

Now, those of us who are familiar with the history of the Iowa Testing Programs believe there are many reasons Iowa has exceptionally good public and private schools. However, we would like to think one is the feedback received in the last 50-plus years from Iowa tests. I'm also sure many people in Iowa would be somewhat insulted by the following statement from the NCEST report:

"In the absence of well-defined and demanding standards, education in the United States has gravitated toward de facto national minimum expectations. Except for students who are planning to attend selective," I'm going to assume that means one of the coasts, "4-year colleges, current education standards focus on low-level reading and arithmetic skills and on small amounts of factual material in other content areas."

And "New forms of assessments—tests worth teaching to—are envisioned."

At this point I would like to specifically address my remarks about testing to the NCEST report, "Raising Standards for American Education." I will also be contrasting the NCEST report to a report of the Office of Technology Assessment, "Testing in American Schools, Asking the Right Questions."

In the authorization for NCEST it explicitly states that "A discussion of the validity, reliability, fairness, and costs of implementing a system of voluntary national tests or examination shall also be included in such a report."

The NCEST report essentially ignores this charge by focusing attention on the innovative nature of the new break-the-mold assessments would make up the national tests. Incidentally, these break-the-mold assessments will probably be very similar to those used in the 1929 Iowa high school program. One gets the feeling from reading the report that these new assessments are so authentic and worth teaching to that issues of reliability, validity, fairness, and adverse impact take care of themselves.

The following rather cavalier statement from the NCEST report exemplifies this attitude:

"Important technical difficulties confront those developing such a new system of assessments. The Council deliberated on these complexities and recommends that special precautions be taken in the development process. First, any system must honor the traditions of local and State responsibility for education and, consequently, must provide flexibility and room for local adaptation.

Second, there are difficulties in producing assessments of high technical quality and fairness. Third, acknowledging that an assessment system of the scope imagined is a new enterprise for the Nation, care must be taken to avoid the unintended and undesired effects of some testing practices, such as narrowing instruction and excluding certain students from assessments. "It will be technically difficult but essential to ensure that new assessments are valid, reliable, and fair."

Now, contrast that statement to the following one taken from the same NCEST report, but buried in the appendix where few will see it:

"Some arguments for a national system focus on principles of good management, incentives, and communication. A set of these arguments implies that our failures with assessment and testing in the past can be overcome with new forms of assessments, assessments more closely linked to teaching and focused on valued accomplishments of students. These arguments are essentially optimistic and have relatively little evidence in their support."

I was astounded when I read the "Report of the Assessment Task Force of NCEST" in the appendix and then contrasted its contents to the body of the report.

Now, consider the following from the OTA report:

"The use of tests as a tool of education policy is fraught with uncertainties. The first responsibility of Congress is to clarify exactly what objectives are attached to the various proposals for national testing, and how instruments will be designed, piloted, and implemented to meet these objectives. The following questions warrant careful attention:"

Then they have a list of 18 questions. I'll mention a couple. "If tests are to be associated with national standards of achievement who will participate in setting these standards?" One Norma would be interested in, "What legal challenges might be raised?"

"If a test or examination system is placed into service at the national level before these important questions are answered, it could easily become a barrier to many of the educational reforms that have been set into motion, and could become the next object of concern and frustration within the American school system." That is the close of that quote from OTA.

I would strongly urge the members of this committee to consider both the NCEST and OTA reports in evaluating the feasibility of a national test. Note especially the reliance on the empirical evidence in the OTA report, as evidenced by extensive footnotes and, correspondingly, the nearly complete lack of documentation in the NCEST report.

I was asked to address NCEST while they were preparing their report to your committee and specifically deal with issues associat-

ed with fairness. Now, I know it may seem strange that a man from Iowa would be asked to address the issue of fairness in a national test, but as one of the authors of a test with Iowa in the title administered to all pupils in such diverse States as Alaska, Arizona, Georgia, Idaho, Texas and Iowa and citywide in places such as Atlanta, Chicago, Dallas, El Paso, and Phoenix, fairness is of constant concern. We long since have removed all references in the tests to corn and hogs.

I can honestly say that I have spent more time in my 25 years with the Iowa program dealing with fairness than any other issue. This includes a number of referred journal publications on this and associated topics in my role as a college professor.

It seemed to me, in my reading of the NCEST report, that issues associated with fairness were addressed in two ways. the first was simply to say that whoever builds these new tests must take steps to be sure that they are fair to everyone. This was the tenor of the earlier quote from the report. As someone who has struggled with such issues for my entire professional life, "It ain't that simple."

The second was to assert that these tests will be based on new methodology, and because of its inherent curricular and instructional validity, such issues are taken care of in advance. It was also frequently stated that little is known about the fairness of these new performance-based methods. I disagree. For this testimony I will only give one citation, but I can add many more if requested.

Feinberg recently reported that majority/minority differences on the California State Bar Exam were not ameliorated with the addition of a performance section. In fact, when results were adjusted for reliability in ratings, mean differences were larger than those that were observed for the multiple choice portion of the tests.

There is a large amount of available data on group performance differences on more traditional multiple choice tests versus performance measures. Anyone with any experience in testing is not surprised by the fact that differences in minority and majority group performance are, more often than not, larger on performance measures. I was surprised at the omission of this literature from the NCEST report.

There are many other issues I could have raised had I had time regarding the NCEST's report. I was especially appalled by the way in which a new council, the National Education Standards and Assessment Council, NESAC, would be appointed. It truly seems that every effort is being made to ensure that anyone who knows anything about testing, and to some extent education in general, will not be a member of NESAC.

I would like to note in closing that creating a new national assessment system will not solve the problems of American education. In fact, it is more likely that such a system would divert attention away from the real problems. I have become convinced over the last few years that from a policy perspective, tests are seen by many as a tool for giving the appearance of solving education's problems without spending any money—that is, without actually doing anything.

The building of a new national test has been likened to the U.S. effort to put a man on the moon. First of all, billions of dollars were spent on space exploration over a long period of time before

any manned efforts were attempted. I would also note that when the effort was launched no one went to the makers of cars and household appliances to find out how to do it. They went to people who knew something about rockets.

I'm not saying that test development is as complex as building the space shuttle. However, like everything else, experience is beneficial and possibly critical.

In summary, a national test is not a new idea. Historical examples show why we have never had one in the U.S. Two, the report on which this discussion is based omits important facts and data about the impact of testing and is inconsistent with other national reports on testing in American schools. Three, the potential for adverse impact has been severely underestimated. Four, the proposal for implementation makes adverse impact even more likely because of expertise.

Thank you.

[The prepared statement of H. D. Hoover follows:]

Statement of

H.D. Hoover
Professor and Director
Iowa Basic Skills Testing Program
The University of Iowa

before the

U.S. House Education and Labor Subcommittee

on
Elementary, Secondary, and Vocational Education

March 18, 1992

As the director of a voluntary K-8 testing program, operated as a service to the schools of Iowa by The University of Iowa, I welcome the opportunity to address this committee regarding the appropriateness and feasibility of a national test. The University of Iowa has furnished tests and services to all schools in Iowa in Grades 3-12 for over 50 years. Currently over 99% of the public and private schools in the state annually participate in the voluntary program I direct using the *Iowa Tests of Basic Skills* in K-8, and a companion program for the high schools using the *Iowa Tests of Educational Development*. Outside of the state of Iowa the "Iowa Tests" are distributed by Riverside Publishing Company, a wholly owned subsidiary of Houghton Mifflin. However, The University of Iowa holds the copyrights to these tests and the university based authors not only operate the state programs in Iowa, but exercise editorial control over the content of the tests and associated materials and services. In the Iowa statewide program schools are contacted once a year regarding participation and they pay their own way. Neither the state department of education nor the state legislature has anything to do with either program. In individual schools results are shared only with that school. All of the high school testing occurs in the fall and the majority of the K-8 testing also takes place in either the fall or midyear. We are convinced that the voluntary nature of these programs is why Iowa is the only state having extensive longitudinal information on the performance of its students, as evidenced by the two CBO reports on achievement trends completed in the late 1980's, following the release of the report "A Nation at Risk" in 1983. The most recent update of these trends is appended to my testimony as exhibit A. I have also appended as exhibit B a table showing how the state of Iowa compares to the rest of the nation in achievement based on data from the Iowa Testing Programs, and independently, the National Assessment of Educational Progress (NAEP). The purpose of these exhibits is to show that by using currently available tests we are capable of not only monitoring trends in achievement, but also of making the kinds of comparisons that led to the reporting of NAEP test results state-by-state.

One of the reasons I wanted to mention the Iowa Testing Programs was because of the amazing similarity of the original high school program started in 1929 to a number of the current proposals for a national test, including that of the National Council on Education Standards and Testing (NCEST) in its recently released publication *Raising Standards for American Education*. The original Iowa high school program was an end-of-year subject matter based statewide academic contest employing a number of types of tests, i.e., not all multiple choice. While this program was extremely popular with school administrators, policy makers, and the press, it had many of the negative side effects alluded to in the NCEST report regarding high stakes tests and was finally replaced in 1942 by a new program using the *Iowa Tests of Educational Development* (ITED) with the entire focus of the program on improving instruction for individual students. The ITED emphasizes critical thinking skills and serves as a model for what a high school graduate is expected to know and be able to do. The ITED has also served as a model for a number of other well known tests. In fact, the first forms of the *Tests of General Educational Development* (GED), the *National Merit Qualifying Test*, and the *American College Tests* (ACT) simply borrowed intact tests from the ITED battery to get started.

Those of us who are familiar with the history of the Iowa Testing Programs believe there are many reasons Iowa has exceptionally good public and private schools. However, we would like to think one is the feedback received in the last 50 plus years from the "Iowa" tests.

I'm also sure many people in Iowa would be somewhat insulted by the following statement from the NCEST report:

> "In the absence of well-defined and demanding standards, education in the United States has gravitated toward *de facto* national minimum expectations. Except for students who are planning to attend selective four-year colleges, current education standards focus on low-level reading and arithmetic skills and on small amounts of factual material in other content areas." (NCEST, p.2)

and

"New forms of assessments–tests worth teaching to—are envisioned." (NCEST, p.6)

At this point I would like to specifically address my remarks about testing to the NCEST report, *Raising Standards for American Education*. I will also be contrasting the NCEST report to a report of the Office of Technology Assessment, *Testing in American Schools, Asking the Right Questions*.

I found the contrast between these two reports very informative. In the authorization for NCEST it explicitly states that "A discussion of the validity, reliability, fairness, and costs of implementing a system of voluntary national tests or examinations shall also be included in such a report."

The NCEST report essentially ignores this charge by focusing attention on the innovative nature of the new "break-the-mold" assessments that would make up the national tests. Incidentally, these "break-the-mold" assessments will probably be very similar to those used in the 1929 Iowa high school program. One gets the feeling from reading the report that these new assessments are so authentic and worth teaching to that issues of reliability, validity, fairness, and adverse impact take care of themselves.

The following rather cavalier statement, from the NCEST report exemplifies this attitude:

"Important technical difficulties confront those developing such a new system of assessments. The Council deliberated on these complexities and recommends that special precautions be taken in the development process. First, any system must honor the traditions of local and state responsibility for education and, consequently, must provide flexibility and room for local adaptation. Second, there are difficulties in producing assessments of high technical quality and fairness. Third, acknowledging that an assessment system of the scope imagined is a new enterprise for the Nation, care must be taken to avoid the unintended and undesired effects of some testing practices, such as narrowing instruction and excluding certain students from assessments. Sufficient safeguards must be built into the system to protect students from negative consequences while the system of assessments is being refined, especially for students who have not been well served by testing in the past.

It will be technically difficult but essential to ensure that new assessments are valid, reliable, and fair." (NCEST, p. 28)

Contrast this statement to the following one taken from the same NCEST report, but buried in the appendix where few will see it.

"Some arguments for a national system focus on principles of good management, incentives, and communication. A set of these arguments implies that our failures with assessment and testing in the past can be overcome with new forms of assessments, assessments more closely linked to teaching and focused on valued accomplishments of students. These arguments are essentially optimistic and have relatively little evidence in their support." (NCEST, p. F-12)

In fact, I was astounded when I read the *Report of the Assessment Task Force* of NCEST in the appendix and then contrasted its contents to the body of the report.

In contrast to the above statements from NCEST consider the following from the OTA report.

"The use of tests as a tool of education policy is fraught with uncertainties. The first responsibility of Congress is to clarify exactly what objectives are attached to the various proposals for national testing, and how instruments will be designed, piloted, and implemented to meet these objectives. The following questions warrant careful attention:

- If tests are to be somehow associated with national standards of achievement, who will participate in setting these standards? Will the content and grading standards be visible or invisible? Will the examination questions be kept secret or will they be disclosed after the test?
- If the objective of the test is motivational, i.e., to induce students and teachers to work harder, then the test is likely to be high stakes. What will happen to students who score low? What resources will be provided for students who do not test well? What inferences will be made about students, teachers, and schools on the basis of test results? What additional factors will be considered in explaining test score differences? Finally, will the tests focus the attention of students and teachers on broad domains of knowledge, as desired, or on narrower subsets of knowledge covered by the tests, as often happens?
- If the Nation is interested in using tests to improve the qualifications of the American work force, how will valuable nonacademic skills be assessed? What should be the balance of emphasis between basic skill mastery and higher order thinking skills?
- If there is impatience to produce a test quickly, it is likely to result in a paper-and-pencil machine-scorable tests. What signal will this give to schools concerning the need to teach all students broader communication and problem-solving skills?
- What effects will national tests have on current State and local efforts to develop alternative assessment methods and to align their tests more closely with local educational goals?
- Would the national examinations be administered at a single setting or whenever students feel they are ready?
- Would students have a chance to retake an examination to do better?
- Would the tests be administered to samples of students or all students?
- At what ages would students be tested?
- What legal challenges might be raised?

If a test or examination system is placed into service at the national level before these important questions are answered, it could easily become a barrier to many of the educational reforms that have been set into motion, and could become the next object of concern and frustration within the American school system." (OTA, p. 27)

I would strongly urge that the members of this committee consider both the NCEST and OTA reports in evaluating the feasibility of a national test. Note especially the reliance on empirical evidence in the OTA report, as evidenced by extensive footnotes, and correspondingly the nearly complete lack of documentation in the NCEST report.

I was asked to address NCEST while they were preparing their report to your committee and specifically deal with issues associated with fairness. I know it may seem strange that a man from Iowa would be asked to address the issue of fairness in a national test. But as one of the authors of a test with Iowa in the title administered to all pupils in such diverse states as Alaska, Arizona, Georgia, Idaho, Texas and Iowa, and citywide in places such as Atlanta, Chicago, Dallas, El Paso, and Phoenix, fairness is of constant concern. We long since have removed all references in the tests to corn and hogs.

I can honestly say that I have spent more time in my 25 years with the Iowa program dealing with fairness than any other issue. This includes a number of refereed journal publications on this and associated topics in my role as a college professor.

The research program of Iowa Testing, related to fairness issues, has been long lasting and very broadly based as Dean Edward Meyen a member of NCEST can attest. His 1968 doctoral dissertation on the feasibility of testing mentally handicapp ▲ students, was supervised by my predecessor, A.N. Hieronymus and titled "An Investigation of Age-Placement, Difficulty, and Importance of Basic Skills in the Curriculum for Educating Mentally Retarded Students". Professor Hieronymus also directed the 1967 dissertation by Beatrize Santos on "Special Achievement Testing Needs of the Educationally Disadvantaged".

It seemed to me, in my reading of the NCEST report, that issues associated with fairness were addressed in two ways. The first was simply to say that whoever builds these new tests must take steps to be sure that they are fair to everyone. This was the tenor of the earlier quote from the report. As someone who has struggled with such issues for my entire professional life, "it ain't that simple". The second was to assert that these tests will be based on new methodology, and that because of its inherent curricular and instructional validity, such issues are taken care of in advance. It was also frequently stated that little is known about the fairness of these new performance based methods. I disagree. For the purpose of this testimony I will only give one citation, but I can add many more if requested.

Feinberg (1990) reported that "Majority/minority differences on the California State Bar Exam were not ameliorated with the addition of a performance section. In fact, when results were adjusted for unreliability in ratings, mean differences were larger than those that were observed for the multiple choice portion of the tests?

There is a *large amount* of available data on group performance differences on more traditional multiple choice tests versus performance measures. Anyone with any experience in testing is *not* surprised by the fact that differences in minority and majority group performance are, more often than not, larger on performance measures. I was surprised at the omission of this literature from the NCEST report.

There are many other issues I could have raised had I had time regarding the NCEST report. I was especially appalled by the way in which a new council, the National Education Standards and Assessment Council (NESAC) would be appointed. It truly seems that every effort is being made to insure that anyone who knows anything about testing, and to some extent education in general, will *not* be a member of NESAC.

I would like to note in closing that creating a "new" national assessment system will not solve the problems of American education. In fact, it is more likely that such a system would divert attention away from the real problems. I have become convinced over the last few years that from a policy perspective, tests are seen by many as a tool for giving the appearance of solving educations problems without spending any money, i.e., without actually doing anything.

The building of a new national test has been likened to the U.S. effort to put a man on the moon. First of all, billions of dollars were spent on space exploration over a long period of time before any manned efforts were attempted. I would also note that when the effort was launched no one went to the makers of cars and household appliances to find out how to do it. They went to the people who knew something about rockets.

I'm not saying that test development is as complex as building the space shuttle. However, like everything else, experience is beneficial, and possibly critical.

In summary:

- A national test is not a new idea. Historical examples show why we've never had one in the U.S.
- The report (NCEST) on which this discussion is based omits important facts and data about the impact of testing and is inconsistent with other national reports on testing in American schools
- The potential for adverse impact has been severely underestimated
- The proposal for implementation makes adverse impact even more likely, because of lack of expertise.

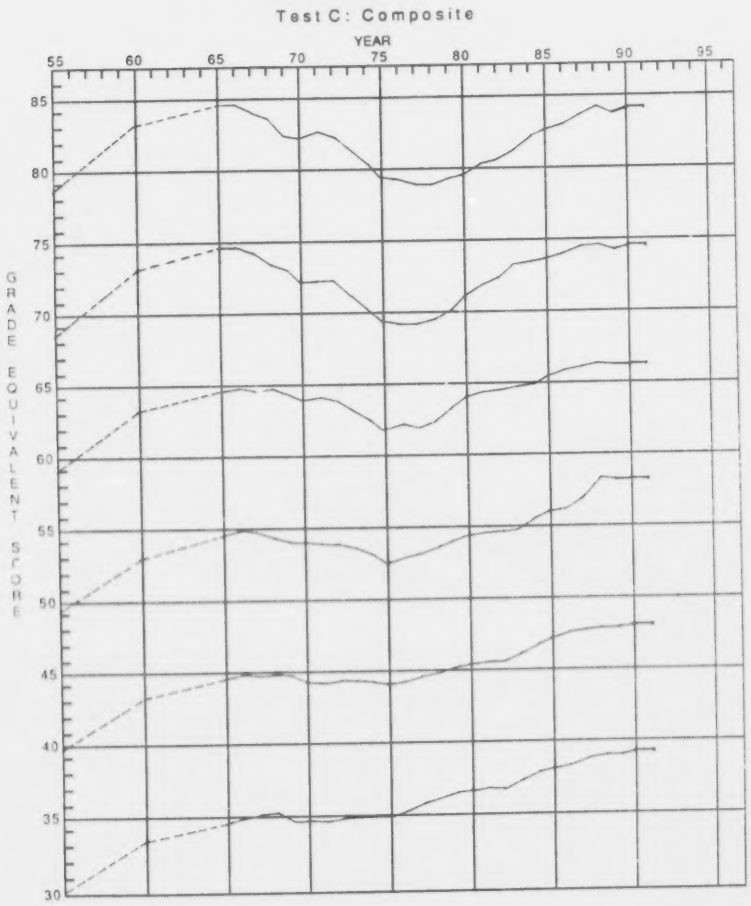
References

Feinberg, L. (1990). Multiple choice and its critics. *The College Board Review*, No. 157.

Office of Technology Assessment, Congress of the United States (1992). *Testing in American Schools: Asking the right questions.*

The National Council on Education Standards and Testing (1992). *Raising Standards for American Education.* A Report to Congress, the Secretary of Education, the National Education Goals Panel, and the American People.

Exhibit A
Summary of Achievement Trends in IOWA, Grades 3-8
in Terms of 1965 "Base-Year" Iowa GE's

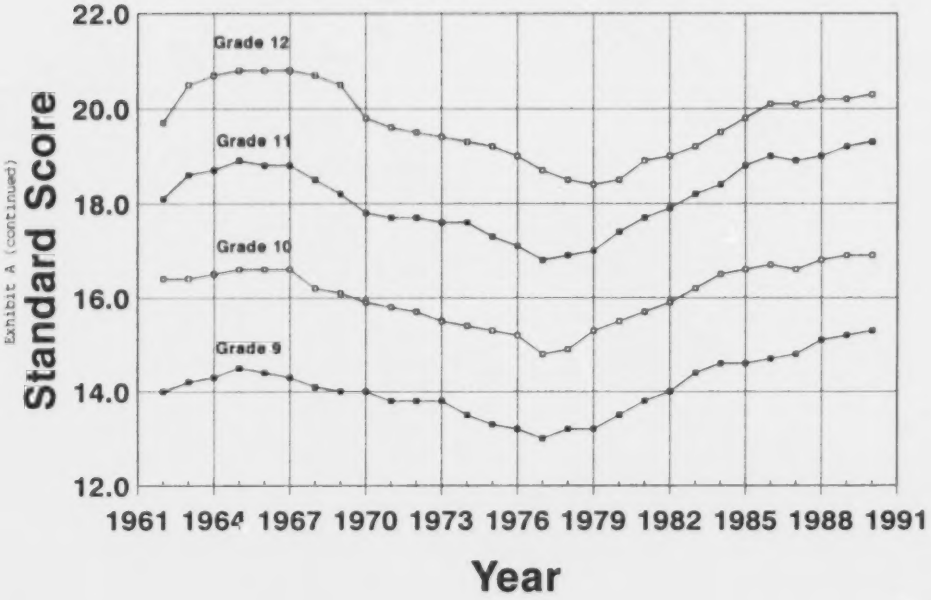


248

Exhibit A (continued)

ITED Average Composite Scores

State of Iowa: 1962-1990



253

254

Exhibit B

A Comparison of the performance of Iowa students to the rest of the nation based on data from the Iowa Testing Programs [Iowa Tests of Basic Skills (ITBS) and Iowa Tests of Educational Development (ITED)] and the National Assessment of Educational Progress (NAEP)

Percent of Iowa Students Above National Median

ITBS		NAEP
Concepts	69	
Problem Solving	63	
Computation	63	
Math Total	69	72
ITED Test Q	62	

Percent of Studetns Nationally Below Iowa Median

ITBS		NAEP
Concepts	67	
Problem Solving	70	
Computation	62	
Math Total	67	67

Sources: ITBS and ITED results based on unpublished reports from the Iowa Testing Programs, H.D. Hoover and Robert Forsyth and NAEP data from "The State of Mathematics Achievement, NAEP's 1990 Assessment of the Nation and the Trial Assessment of the States." National Center for Education Statistics

Chairman KILDEE. Thank you very much. I want to thank the panel.

To Dr. Rezmierski, I attended school at St. John's Seminary near your place there and also had a number of friends at St. Mary's Orchard Lake, from which I picked up some of my Polish so I was happy to be able to respond to you some in that great language.

Let me ask you this question. The Council recommends that content standards be developed on a national level but that school delivery standards be developed on a State level. School delivery standards are a way to determine if the child failed the system or whether the system failed the child. Do you believe that schools should be held accountable just as the council recommends that students be held accountable? Do you want to go down the line and respond to that?

Ms. Cantu?

Ms. CANTU. Actually, yes, I do have a strong opinion about that, that school delivery standards should reach down to the school level. A good analogy is the Federal Equal Education Opportunity Act that protects the rights of limited-English-proficient students. That Act applies both to the State education agency and the local education agency. We have used it to protect the rights of students against individual districts. The most recent example is the Denver Public Schools. In that situation, we didn't sue the entire State of Colorado, and we were able to work out a settlement with the Denver Public Schools.

The model of accountability on the district level is essential because there are a number of States that are decentralized in the delivery of education services, and there is diversity within States. Either through lack of resources or lack of will some districts do much worse in serving their students than others. Actually, what I would recommend would be both accountability at the State level and the district level.

Chairman KILDEE. Dr. Hoover, would you like to respond to that?

Mr. HOOVER. Yes, I agree. I think, though, this brings up what is really the serious issue here and what originally gets to the reason that I focused on tests, is because the national content standards mean that ultimately we must be talking about national performance standards. Again, it's noted that ultimately all of these assessments have to present, produce what we call comparable results in attainment of standards.

Now, given that, the opportunity of students to learn, which is the fundamental issue I think I tried to point out at the end, is that the major problems in American education deal with those delivery standards. I don't see how this report addresses any of that.

Chairman KILDEE. Thank you.

Dr. Rezmierski?

Dr. REZMIERSKI. Mr. Kildee, I would support both at both levels, as I can make some interesting comparisons. In our State, in Michigan, we have PA 25, which I'm sure you're aware of, that forces us who look at ourselves as a system before we simply assume the child is wrong, the student is wrong. To quote Dr. Spader, "The parents don't send us their second best; we shouldn't be giving our second best."

That means we have to look internally to ourselves, whether it's staff or process delivery. It oftentimes is not the child. Setting these standards has to be a whole package where again we focus, like Walt Disney World does, around the client: What is that client as an ultimate outcome going to be doing? Can they demonstrate that, and, if they can't, where has out system failed that learner?—not assuming the child didn't learn it in a lock-step time when we were ready to teach it.

Chairman KILDEE. One thing I discussed informally with Mr. Kearns this morning, pardon me, was at the 4th grade level we probably have lower risks. At the 12th grade level we might have more high risk, according to some recommendations anyway. But suppose, let me just give this scenario to you. Say, at the 4th grade level we were to have standards and then testing or assessment to see whether the students have achieved those standards.

Would it not be appropriate, then, that if we found out—I can predict, as I mentioned before, in certain schools that students are going to do better than in other schools. I can go out to Fairfax County here, where my children went to school, and predict that they generally will do better because of good family support, a variety of reasons, than maybe some other school system.

Suppose we had standards in the 4th grade and assessment in the 4th grade. Would it be appropriate then where we found that the students were not reaching the standards that we have a combination of local, State, and Federal programs, dollars really, to help intervene at that point for remediation? I'm always afraid of just gathering statistics with tests and not having that used as a tool to improve education. Would any of you care to comment on that?

Ms. Cantu?

Ms. CANTU. I'm unclear in terms of how the intervention would—consist of, I mean. For example, if the intervention is a negative one where funds would be withheld because people weren't satisfied, I think that's counterproductive. But if it's a positive intervention in terms of being an advocate for that particular school or that system, I could see it having a very good effect in terms of turning that school around. Unfortunately, though, you really just can't take a snapshot of the educational process and just pick the one grade. But it's a start. It absolutely is a start.

Chairman KILDEE. What I was thinking of is possibly some type of remedial programs maybe using the Chapter 1 and, say, those schools where assessment indicate that those in the 4th grade are not meeting the standards, that we give some flexibility to the Chapter 1 and maybe some more dollars. Where there needs to be some greater effort to help these students achieve and measure up to the standards. Would that make any sense? Does anyone want to——

Dr. Rezmierski? You're a superintendent.

Dr. REZMIERSKI. Mr. Kildee, yes. I would assume and agree that that would be a step in the right direction, but I think again part of the problem that I'm having with the discussion in the whole area, not your question in particular, but it is that we are constantly reinventing the schools that we went to.

We have to really look at what kids need to know in the 21st century and focus on that. If we continue to lock-step everything in and let the calendar drive our approach to education, we will forever repeat the historical mistakes we have envisioned and seen over the last 100 years in education, not all wrong. We have obviously moved a great Nation forward and have accomplished many positive things.

As we look at what our students need, that future learner, I think we need to reexamine that whole structure. We know, for example, in relationship to your question, that Head Start has proven successful.

We have even in our regional area many examples of many successful young people because they got an earlier start, because we were able to excite them, redirect them and, in effect, galvanize the teachers to reapproach learning for those students because they had failed; not as, "You failed, too bad, let's go to the next chapter," but, "Now, what is it that you didn't learn that we consider essential learning, and let's go back and reteach it."

Again, maybe the fault was ours. We taught it in a modality that wasn't acceptable to that student or possible for them to learn. I think it's a good start, but I think there is a bigger question out there about the restructuring of our Nation's schools that's at the heart of much consternation.

Chairman KILDEE. Let me direct this question to Dr. Hoover. It's related to some things you said. Are U.S. students as poorly educated as many reports appearing in the press would indicate? Many news accounts suggest that American students are performing far worse than their international counterparts. Could you enlighten us on that?

Mr. HOOVER. Well, first of all, I think that much of what we read in the popular press about this is just flat wrong. You should read the papers by R. S. Rockberg and Jerry Bracey that have been in the press in the last year or two. There are things about those reports that I disagree with that indicate that there are all kinds of major problems in making those kinds of comparisons.

The latest stuff, and as a great example of what standard setting can do in coming out of NAEP, in which they say that, gee, only 5 percent of the kids in college—or 5 percent of the 12th graders are capable of doing college work. It's a classic example of someone taking statistical data, creating a standard-setting process that is dramatically misleading.

When you look at the NAEP data, take that very same data that says only 5 percent of the kids reach the standard of advanced. But then, for example, in the State of Iowa you take the items that exemplify advanced standing, on the average 65 percent of the kids in Iowa gets all of those items right.

Well now, which is right? As far as I'm concerned it would indicate to me that 65 percent of the kids in Iowa—and nationally it's very similar figures—are capable of doing college work. The international comparisons, again, become very cloudy because, for example, in a country like Switzerland, that I'm fairly familiar with for a couple of reasons, they track kids off at age 11. They have a wonderful school system, but they use tests in ways that I think all of us here would find somewhat abhorrent.

Also, related to this in terms of the points that Norma was making about is resources and trained personnel. It happens that my son has a, I guess I would say, fiancee who is a school teacher in Switzerland, who is a second year teacher, has the equivalent of what we would call a normal degree in the United States, teaches at the vocational level after these kids have been tracked, so this is a junior high level teacher. Her salary is more in U.S. dollars than a friend of mine in Iowa City who has taught 25 years and has a Ph.D. in the public schools. Those are incredible differences in both resources and your ability to get trained personnel.

Now, I'm not making the indication—math and science performance in the U.S. is pretty abysmal. I'm not going to disagree with that. But none of these proposals that I see will have any impact on what the real problems are with math and science performance in terms of better training of teachers and the more demanding, in a sense, standards in math and science. I don't see how these will lead to this.

Chairman KILDEE. I'll probably come back for another round of questions.

Mr. Reed?

Mr. REED. Thank you, Mr. Chairman.

Dr. Hoover, I was intrigued by your testimony in which you indicated that—and I want to be accurate so correct me if I'm inaccurate—that minority students do better on multiple choice tests than on performance tests. First, is that an accurate recitation of your testimony?

Mr. HOOVER. Well now, the differences in majority and minority performance, the differences between those two groups and how well they perform, those differences in general, now not always, but in general, tend to be larger on performance assessments than they are on more traditional multiple choice tests.

Mr. REED. Again, so I fully understand you. If you compared majority students, minority students multiple choice scores, there is a certain gap?

Mr. HOOVER. Yes.

Mr. REED. If you looked at performance testing, there is a larger gap?

Mr. HOOVER. In most cases.

Mr. REED. In most cases.

Again, I am asking these to try to understand what is going on. I don't have any sort of conclusion, maybe you do. But there is a sense I had previously that one of the reasons that we don't want to use multiple choice testing is because it is an inaccurate assessment of minority performance. In fact, I always sense a subtle sort of push for more performance-oriented tested because that's a better predictor of actual—well, whatever you're actually trying to test. That's one comment.

A second comment is that the other sort of premise behind some of these performance testing is that it replicates much more closely what actually people are responsible for either educationally or professionally, et cetera. I am just concerned, intrigued, et cetera, about what this all means.

52-840 O - 92 - 9

Having opened up this large can of worms, would you like to provide some further background—and your colleagues might want to comment also?

Mr. HOOVER. Well, first of all, I'm a very strong proponent of performance testing. That may not have come through here because of the position I've been sort of put in here. Because one of the problems when you build tests like I do, they have been overused, used for things they shouldn't be used. We really do need a much broader base of tests. I strongly believe that.

But when it comes to issues associated with bias, first of all, since multiple choice tests like the ITBS or the SAT or the ACT are what people have been shooting at for a long time, that's mostly what you hear about when you say tests are biased. Those are the ones that are being used, so they are the ones that are going to be criticized; and they should be. But because of all of this intense scrutiny that these tests have had over the years, we do our damnedest to make these tests as fair as possible.

Now, with multiple choice tests, with tests where you can ask lots of questions, you can at least start to balance out things like regional differences. In fact, I'm criticized much more frequently for test items that will favor a student, let's say, in Arizona versus one in Minnesota than test items that will favor a black student over a white student.

But if you ask lots of questions, you can broaden your sample and deal with the kinds of issues—the context, the items in—that will be unfair to individual students. When you have a more limited number of tasks, which is invariably going to be the case with performance assessments unless you get a lot more time, you increase the chance that you're going to ask things that are going to be unfamiliar and create situations that will be unfamiliar to more students. It just sort of has to happen that way.

Mr. REED. Let me ask, again another question which is abstract perhaps. There always seems to be dilemma at the heart of any of these situations of testing. It's the pull between a fair test, whatever that means—and it's a very difficult thing to define—and a valid test, one that actually serves the end that you want.

I mean, an example is that the simpler you make the test, the more people will pass the test. You eliminate the fairness question because everyone does great and no one complains. The harder you make the test—and again we're talking about relative things and I'm using rather imprecise language—the more there is pressure to make that test fair.

I'm just wondering—again this is a large philosophical question as well as a policy question—what type of comment you might have in terms of that dilemma. Because essentially that's what we're struggling with here on this one aspect of this issue, and we have to have, I think, a perspective and a basis to make a good attempt at that. You are an expert and your colleagues are experts, so your comments would be useful.

Mr. HOOVER. Well, I think you can have very demanding tests that are fair. I really do. This is one of the problems, though, with I think the current trying to—well, I mean, when I'm in Washington I'm always amazed. I hear the word "consensus" about 80 times.

If we try to build tests that are going to make, that are going to be under this sort of "top down," these have to be fair in all these ways, where everybody gets their shots in at them—and primarily the shots will be done by people who probably don't know much about testing—we're going to eventually end up with the lowest common denominator.

This is what has happened in virtually every State-mandated minimum competency test. I mean, this is the primary reason I am opposed to the national test. I think ultimately, regardless of how wonderful it sounds that we're going to have these high standards, that when we get to the end, the test, which is the linchpin of all this, is bound to be not very good.

The test I mentioned here, the Iowa Test of Educational Development, that we use in our State program—we use it in our State program because it's not adopted in any way—is very seldom used outside of Iowa anymore. The reason it isn't? It scares the hell out of people!

Mr. REED. Why, Doctor?

Mr. HOOVER. It's too hard. They look at it and they say, "God! My kids will do poorly on that test." That is fundamentally the reason. It measures critical thinking and all of these wonderful things. And as I said, I'll stand up to that, I'll let you go through that test and you will be convinced. But it gets very difficult then to get a test like that through this system.

Mr. REED. Essentially, some of your criticism is the fact that because of the forces, the political forces, and I say this with a small "p" not the things we're witnessing today or yesterday but the small "p" politics is that you're not going to get that kind of rigorous, valid and, hopefully, fair test adopted; you will get the least common denominator?

Mr. HOOVER. That's pretty much my view, yeah.

Mr. REED. Let me——

Ms. CANTU. Well, let me add something to that.

Mr. REED. Yes, ma'am?

Ms. CANTU. The fear of the test is more the fear of the consequences of how the test is going to be used, okay, so it is not a fear of the test. In the Hispanic community, we have high expectations of the students, but we have a fear of how the test is going to be used. That's why I gave the example of the bilingual teachers in California. They are afraid that if they don't learn enough Spanish so they can communicate to the children in their classroom, they are afraid they are going to lose their job. That's why the waivers are put into place.

They are not afraid of the test. They are going to take that test that demonstrates their ability to speak English and Spanish as many times as you want them to take it. They are afraid of the consequences of what happens after they fail it a certain number of times.

The response, the political response, is they get a waiver for 7 years. Now, the students are tested, but they don't get waivers. If they fail, they get held back in grade; and all the studies show that a child held 2 years in a grade is a dropout. That's a child at risk, and that's a dropout. They have a consequence that they can't escape.

It is the fear of the consequences that's troubling us in the testimony today. It's not higher expectations, it's not having higher standards, it's not having a system of assessment. It is, once you create this, what consequences are we back in all the parts of the United States going to face because we created this?

Mr. REED. But isn't there a real concern, or I think there is, of the system in which there are no consequences? I mean, you've just recited an example of people who take tests repeatedly to demonstrate their fluency in Spanish or another language, in English, fail those tests and still teach.

Ms. CANTU. That's right, and the consequences are borne by the children——

Mr. REED. Well——

Ms. CANTU. [continuing] not by the teachers, not by the State of California. You don't have the State of California superintendent in front of you today and you're not talking to that person. The consequences fell on the children in California.

Mr. REED. But don't the children of California bear the consequences of repeatedly spending 8 years in school and not being able to read and then getting either a diploma or a certificate of completion and going into the world?

I mean, that's why I'm not suggesting that we have any sort of definitive answer at this point, but one of the problems is, you don't want a world in which the consequences are unjust, inequitable, and irrational, but I don't think you want a world, too, in which there are no consequences after 8 or 12 years of education.

Ms. CANTU. Unfortunately because of politics, the consequences fall on the more insular populations. The limited-English-proficient children who can't be hear and speak up for themselves bear consequences. They don't have—they are not empowered, and they can't be here representing themselves so they bear the consequences.

The teachers who graduate from the universities and who do receive the diplomas but aren't prepared to teach the language-minority children out there, those teachers don't have a consequence on them. It ultimately falls on the least powerful group.

Mr. REED. Well, I see. I agree with you. But what I see is that these least powerful groups are bearing the consequences now without an assessment regime. I don't think anyone is standing up and saying they're doing terrific and they're really involved and they are on their way to success.

What we are grappling with is trying to, first, see if we can develop a system in which you assess and evaluate. I think all of us here on this panel have severe questions about how you do that, should you do that. Maybe it's better to refocus, as I suggested with the Secretary, into helping teachers be better teachers rather than assessing children.

Have we created a system in which there are no consequences, where everyone sort of goes through the system, gets out having spent 12 years without 12 years worth of education?

Ms. CANTU. We have consequences. In Texas—and this goes to the point Mr. Sawyer made about incremental consequences. In Texas, 4 years ago, we put in an exit test requirement that every high school senior has to pass this examination.

Okay. The first year, everybody passed it. The second year, everybody passed it. The third year, a few people failed it. We're in the fourth year now, and the failure rate is almost 60 percent failure rate. We took a huge jump between the third year and the fourth year. People are screaming foul.

What happened here? It was an incremental. They had 4 years to get ready for it. People were 8th graders when this idea was approved, and now they are screaming foul, "I'm paying consequences, I didn't know these were going to happen." People are afraid of situations like that.

Mr. REED. Well, but it goes back, I think, to what Dr. Hoover said, is that he is not afraid of having a tough, hard test like the Iowa Test go out there with high expectations and validity, et cetera. That might be or might not be what is taking place in Texas. What you're saying, then, is that that's not what you want?

Ms. CANTU. What I'm saying, it was a hard test to begin with. Politicians changed the cut score every year, every year, every year, every year——

Mr. REED. Okay.

Ms. CANTU. [continuing] but didn't change the curriculum, didn't change the level of resources going into the schools, didn't change how teachers are being prepared, didn't change the respect given to teachers, didn't change parental involvement, didn't change the classroom sizes. All of those variables of quality remained the same.

Mr. REED. Are you using now these results as a rallying point to do those things? You know, is it even a sort of unintended effect of the test that now you can point with a great deal of clarity and precision to the lack of resources, the lack of teachers? Is that an effect, or am I being too optimistic?

Ms. CANTU. You're being optimistic. People are blaming the children.

Mr. REED. Okay, all right.

Ms. CANTU. You're being optimistic.

Mr. REED. Okay.

Ms. CANTU. I mean, they are not blaming the system at all. They are saying, "Those children failed."

Mr. HOOVER. That's right.

Ms. REED. I thank you because they are very insightful comments, and I appreciate them.

But just one final general question. It goes back now to kind of this notion of international competition. Because a lot of the discussion about why we should test is economic competition—"everyone else does it." I think, Dr. Hoover, you have had a lot of experience looking at other international testing schemes. I asked the Secretary about this too. What is your sense of our position vis-a-vis the other nations in terms of standards and testing?

Let me go back and add one final point is that you're sort of saying, well, you don't want a national test, et cetera. But it seems because of the nature, the structural nature, of other countries, they have national systems and national tests. Having wandered into this question, would you infuse it with some insights?

Mr. HOOVER. Well, a little bit, a lot of this business issue, the fact that we supposedly can't compete businesswise, and it's interesting

to hear Secretary Kearns who was a CEO of a big company or to hear someone from someplace like Ford or General Motors talk about this. Because, frankly, now this is an opinion of somebody who is not an economist or anything, but me——

Mr. REED. It adds more credibility.

Mr. HOOVER. [continuing] just a boy from Iowa, it sort of looks to me like the real problems we face are not because the people working on those assembly lines aren't capable of doing what they can do and that they are poorly educated. All of the bad mistakes that we have made, the reason we can't compete now internationally is because of graduates of Harvard and Yale Law—Business School and I might add the University of Iowa, not the law schools——

Mr. REED. Thank you. I graduated from the Law School, so the business school, I blame the business school.

[Laughter.]

Mr. HOOVER. [continuing] that, in fact, have made some incredibly bad decisions and find the schools fairly simple scapegoats for all this. Now, we supposedly have the best university system in the world. Now, you know I'm a product of a one-room country school in the Ozarks and two good public universities, I think, the University of Missouri and the University of Iowa. I really get sort of steamed about people from Exeter and Yale talking about the public schools. They wouldn't know a public school student if they came up and bit them on the leg.

[Laughter.]

Mr. REED. We don't have any problems with that.

Mr. HOOVER. Also, since I'm rolling here, my son is a graduate of one of these schools, one of those dumb high schools in Iowa, and a student at Iowa State University and will graduate this spring. He spent a whole year in Switzerland at EPFL, sort of their MIT. He didn't have any problem competing with those people, none at all, and he even had to do it in French.

I mean, we are finding the schools convenient scapegoats. We have wonderful schools in the United States. We have some incredibly big social problems we're not dealing with very well. None of this stuff is going to solve any of those problems. I am very serious, it's going to end up diverting attention from what the real problems are. We blame the kids. As Norma says, that's the easy way, to blame the kids. Again, internationally, I can't believe we win all these Nobel Prizes with all these stupid people.

Mr. REED. Well, thank you for your diplomatic response to my question. I guess I have to apologize because I spent a couple of years at Harvard, as I think Norma did, too, so maybe we're the exceptions to the rule. But I always credit West Point—a good public school—where I stood at attention for 4 years and learned the value of good posture.

Thank you very much.

[Laughter.]

Chairman KILDEE. Thank you very much.

Before we go on to Mr. Sawyer, Dr. Hoover, you offered to give the committee the Iowa Test of Educational Development. Would that be a high-risk test, do you think, if we took that?

Mr. HOOVER. Well, it would if we published in USA Today, the results.

[Laughter.]

Chairman KILDEE. Mr. Sawyer?

Ms. CANTU. It's that fear of consequences again.

Mr. SAWYER. I particularly appreciate Ms. Cantu's last set of remarks. The fear of the test is really the fear of the consequences. When I asked Mr. Kearns about those things that would enable us to achieve simultaneity instead of some measure of sequential development in getting from here to there, he spoke about the importance of beginning with low stakes.

It seems to me that that is a very clear and perhaps unintended but nonetheless candid expression of the fear of consequences, the whole inability to get around the Lake Woebegone approach to testing, and the reason Iowa is such an unpopular test these days.

I just have a couple of observations. I really don't have questions for you, except to say thank you to all of you.

You spoke about the analogy of taking a temperature. Taking a temperature is no problem, but if it's not really functioning as a diagnostic tool, if it doesn't lead to other arenas of diagnosis and remediation, then it does no good at all. There is no treatment. It is simply diagnosis for it's own sake.

Dr. Rezmierski, I appreciated what you had to say about the importance of full disclosure. It's important to know what's humor and what's grim truth, I think in some cases. I could not help but think when Dr. Hoover was talking about eliminating references to hay and corn in the Iowa test. A few years ago in Ohio, which is often mistaken for Iowa anyway——

Mr. HOOVER. And Idaho.

Mr. SAWYER [continuing] we pointed to the problem in Ohio of the use of the word toboggan, which in the northern part of our State refers to a bentwood sled, and in the southern part of our State, to the kind of knit cap that you wear when you go bentwood sledding.

It reminds me of a true story about the difference in urban areas and rural areas between what is meant by service. I mean it goes back to the punch line of a myriad of very old jokes. In my case, it was the difference between if you're on a farm and you have a sow and you want to go into the pork business, you have that sow serviced. I mean it's a very useful term if you're going to deal in forums like this one and be able to get away with a printed record, such as we're asking our secretary to take down today. In an urban area, it meant something very different.

When I was first in the State legislature, Mr. Chairman, I used to recognize that there were real differences in language between the northern, urban part of Ohio and the southern, rural and agricultural part of Ohio. I never discovered that as completely as one day when coming back to the general assembly in Columbus, I confronted one of my rural colleagues who asked me what I had done over the weekend. I said, well, I had been at home doing constituent service. He looked at me and he said, "Son, given your voting record, I do believe you may be right."

[Laughter.]

Mr. SAWYER. The real problem that we have here is that we may be seeking to provide a kind of service across this country that will

have exactly the same consequences as that rural colleague perceived in what I was doing over that weekend.

It is not too much to ask questions about how we get from here to there. It is not too much to ask questions about the importance of knowing what we're seeking to measure so that we can test those questions, however difficult they may be, of validity and fairness and reliability. It is critically important if we believe in fundamental things, like the notion that what you measure may, in fact, be what you get. It goes back to the point of Dr. Rezmierski's bumper sticker.

It is not simply a matter of being able to go out and develop a measure that in the marketplace will provide a test of children and schools and systems. It becomes a question that I think is the only point in what I've heard you say this morning that I would really disagree with strongly.

I believe that there are some people who would use this system of testing as an instrument of market measurement and who would suggest as a consequence of what we measure that not only are children at fault, but an entire approach to education is at fault and would use the results, the consequences of these tests, to dismember something that needs more fixing than restructuring.

I really worry about that as much as anything I can possibly say today, except to say that a set of goals, if they are allowed to corrupt what we're doing in terms of our educational system, may well become the self-fulfilling prophecy that is the undoing of our educational system.

I worry more that when we set a goal that we ought to be able to achieve supremacy in math and science education by the end of this century; that if we do not, it be used as the excuse for wholesale change in a system of education that has been the model for the rest of the world for a century.

I am more worried about that, than anything we can possibly say here today. That's why it is so critical that we not compromise the way in which we measure our performance. There is so much that depends on it. I'm babbling now, and I'm going to stop. But I thank you all very much for being here. It is an important panel to us, for us today.

Thank you, Mr. Chairman.

Chairman KILDEE. Thank you, Mr. Sawyer.

Two additional questions, and the committee may wish to submit further questions in writing to the three panelists.

To Ms. Cantu: How have standardized exams been used, or perhaps misused, with limited-English-proficiency students? Is there a problem with the standardized testing of those students?

Ms. CANTU. The standardized tests have been used both in a positive and a negative way. The positive way is what we would all optimistically hope tests would be used, in the form of assessment, children who need additional services are identified through the use of testing instruments. The negative way is that they are used to label children, to track children, to put them in low-ability groupings. Until someone is satisfied with an arbitrarily cut score, a child will be labeled as low-ability for years and years.

So there is a little bit of both going on, Chairman. There is some real proactive use of testing as a way of delivering service to every child that can possibly be identified.

By the way, there are already out there many well-respected assessment instruments for identifying the limited-English-proficient children. What is lacking is what happens after you have identified the child? What is the follow-up? Do you have a trained teacher that can communicate with that child? Do you have a small enough classroom setting? Do you have the materials? I mean, all of these are questions where most of the problems arise cross the country. But finding the children, we can clearly do that.

Chairman KILDEE. Now, along that line, I think when we began hearings over a year ago, I gave an example. I taught Latin for 10 years, and in Latin—I got that in the seminary. In teaching Latin, I always understood the sequence of tenses very readily because we had to speak Latin for 4 years in all of our classes, not just our Latin classes. When you went from the indicative mood to the subjunctive mood, you just knew what tense to use.

My students, through testing I found out that even some of my better students just could not understand the sequence of tenses, and I would keep giving the test. Finally, I decided I had to redo the method of presenting that, and I did. Then I reached a point really where my poorer students could really understand the sequence of tenses. At least mechanically they could know which one to transfer to.

There is where testing really made me go back and examine my teaching methods, and that's where I think testing can be very, very effective. But if testing is just gathering statistics of things we probably already know, it isn't really very productive for education. I think that testing which will help bring about some changes in education whether it be the individual teacher like me, as me the Latin teacher, or the system, can be very, very helpful to education.

Dr. Rezmierski, a question of you. You're not only a superintendent of a great school system in Michigan, but you're also with the Council for Exceptional Children. How are exceptional children treated in the current assessment program? Are there any positive experiences from current practices that we should consider?

Mr. REZMIERSKI. Well, first and foremost, Mr. Kildee, clearly, it would be safe to say that most districts try to find ways to avoid having the exceptional child tested so that their scores aren't reflected negatively or downward, as other panel members have indicated in other ways, in other methods. That's one problem we have because we're all caught up in this "Where are we going so fast" notion that it wouldn't surprise me if we soon would be competing with people on Pluto, to be comparative.

We have to be sure we know where we're going, because as we're using these tests not only in the State of Michigan but nationally, we have to know why we're comparing them, because we're placing this pressure point on the teachers and school districts, administrators, boards of education who maybe look at selectively not testing certain kids.

We also must limit our base of delivery, as you've mentioned, in mastering learning, come back and reteaching it as a simple paper-

and-pencil test. A lot of kids can perform, but they can't do it on the paper-and-pencil test in 20 questions on Friday for 30 minutes. We have to expand our capacity to evaluate and validate what all kids can do.

We have found successful entities, the other part of your question, by having the test read by fellow high school students and having it on tape-recorded messages so that parents could be involved. We have found no one of an exceptional student category who has cheated. They tried to do the best they could. With all those other supports—and it didn't dilute the value of the test for all the rest of the kids, nor did it score our district to be considered less. In fact, we're still in the top 10 in the State, even in the revised test in Michigan.

We don't feel we have been sullied by the fact that we're trying to include all kids. The difference is that I think too often they are put to the side and it aids and abets the tracking notion.

Chairman KILDEE. Thank you very much.

I want to thank the panel. I think you have really given us some insight into something we're really wrestling with. We want to do what is good for the education of our students in this country. That's really what it is, to help those students get a better education. If standards and testing can be devised in such a way as to help improve education for the students, then we will have achieved something.

I don't want to be racing down a road where we're merely gathering statistics to know what we already might suspect. That's really not our desire at all. If we can help improve education for our children through standards and testing, then we might achieve something that we will be proud of.

You have been very, very helpful to us this morning. We will leave the record open for 2 additional weeks, and we may submit some additional questions to you in writing. Thank you very much.

At that we will stand adjourned.

[Whereupon, at 12:35 p.m., the subcommittee was adjourned, subject to the call of the Chair.]

[Additional material submitted for the record follows:]

STATEMENT OF DR. ANN LIEBERMAN, PRESIDENT, AERA, TEACHERS COLLEGE, COLUMBIA UNIVERSITY

Mr. Chairmen and members of the subcommittee, we are pleased to provide this written testimony for the record of the oversight hearings about the proposed National Education Standards and Assessment Council. AERA has been vitally interested in the issue of national standards and assessment for several years, and held an Invitational Conference on the topic last June. The papers from that Conference were printed in a special edition of *Phi Delta Kappan*, and Linda Darling-Hammond and I have summarized the position represented through this research in the *Chronicle of Higher Education*.

The governing body of AERA, after a review of the research of the past 20 years, concluded that the preponderance of evidence argued against the use of high-stakes tests to achieve educational accountability. Indeed, we feel that irrevocable harm will be done to children if we proceed to establish the high-stakes national tests being proposed by some policymakers. It is from the perspective of reviewing this research and holding these deliberations that we now approach the proposal to establish a National Education Standards and Assessment Council (NESAC).

We have been attending the oversight hearings on the report of the National Council on Educational Standards and Testing held recently in the U.S. House of Representatives. Much of this testimony echoes our own views. Certainly the testimony garnered by the House Subcommittee on Elementary, Secondary, and Vocational Education buttresses our specific recommendations regarding the provisions for NESAC, which are as follows:

1. NESAC should be given broad responsibility for stimulating the development of national content standards.

Current Senate legislation (S. 2) provides for NESAC to "be a coordinating body to ensure the establishment of national standards." More than a "coordinating body" is required to establish national standards and curriculum frameworks. We need a group that will motivate and stimulate citizens and professionals at the State and local levels to become involved in this important task. NESAC should conduct hearings and disseminate information about procedures and approaches to development of content standards; it should undertake research and evaluation programs related to standards and curricular frameworks.

2. NESAC should be given major responsibility for stimulating establishment of an R&D program related to national assessment questions and for monitoring its performance in meeting the information needs of the Council. Almost all of the witnesses appearing before the subcommittee testified to the need for a research program to investigate the many unknowns related to assessment: new performance assessments, costs, and the effects of assessment on instructional systems and students. In our view, simply calling for "break the mold assessments" independent of a substantial research and development program is irresponsible. Subjecting teachers and students to new assessments or new uses of assessments on faith alone is similarly irresponsible.

The accompanying testimony was prepared for the House Subcommittee on Select Education oversight hearing (March 17, 1992) by the co-directors of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Their testimony outlines an agenda of essential research questions to be explored about national assessment. Informed policy requires that these questions be addressed as an essential component of the move toward national standards and assessment.

3. The membership or procedures of NESAC should be configured to assure that it has direct access to the expertise required to perform its functions. Certainly, there is an important role for practitioners, policymakers, and lay citizens in this work. However, the membership design proposed in the Report likely will result in a body devoid of experts in school reform, curriculum design, cognition, and assessment.

4. NESAC should not be assigned responsibility for certifying standards or assessments. The States and local districts should be the first and final voices heard on the topic of adoption of standards. The NCEST Report expresses the belief that a relatively small group of appointees can represent adequately the diversity of the education public, and that such a group will be "accountable." These do not seem to us to be reasonable assumptions. Moreover, the idea of certification of assessment instruments flies in face of a basic understanding in the testing field: Validity of instruments cannot be separated from their actual use. Tests can not be judged to be valid in general, or in the abstract, or in advance.

We believe that you will find evidence supporting these positions in the recent OTA publication, *Testing in American Schools: Asking the Right Questions*, and in

recent expert testimony provided in oversight hearings of the Subcommittee on Elementary, Secondary, and Vocational Education. Drs. Resnick, Smith, Haney, and Darling-Hammond all indicated that standards and curricula should be developed before assessments (2/3/92). At a subsequent hearing (2/19/92) Drs. Koretz, Feuer, and Kean made similar recommendations. Also, each of these individuals called for a strong R&D program on assessments prior to embarking on a national examination system.

It is significant that Governor Romer, the driving force of the National Goals Panel and the National Council on Educational Standards and Testing, concluded his remarks about the Council before the subcommittee (2/19/92) in the following fashion (we are forced to paraphrase as this was an extemporaneous comment): "It is important that we get a structure in place, but if assessments are a problem we should begin with standards. A strong R&D program working with the State and local governments is a good way to proceed in setting up national assessments. And if the certification role of the Council is a problem for the Congress, I could live without it."

In our view, these expert witnesses and Governor Romer are correct. The Congress should create a national entity which has responsibility for stimulating development of national standards and curriculum frameworks by encouraging broad State and local participation of all the publics concerned with education. The Council membership should be modified to include persons with expertise in areas of curriculum development and assessment. The Council not be charged with creating national assessments for individual students, or with certification of standards or assessments. The Council mandate should be expanded to include stimulation of a research program sufficient to answer policy-related questions such as those posed in the CRESST testimony.

We appreciate this opportunity to address these important questions. If you wish, we will be pleased to provide you with a set of the research papers providing the foundation for our thinking on this topic. Please call on us if we can be of additional assistance.

STATEMENT FOR THE RECORD BY
WILLIAM H. KOLBERG

ON BEHALF OF THE
NATIONAL ALLIANCE OF BUSINESS

BEFORE THE
SUBCOMMITTEE ON ELEMENTARY, SECONDARY, AND VOCATIONAL EDUCATION
COMMITTEE ON EDUCATION AND LABOR
U.S. HOUSE OF REPRESENTATIVES

ON "ADOPTING RECOMMENDATIONS OF THE NATIONAL COUNCIL ON EDUCATION STANDARDS AND TESTING"

March 18, 1992

Mr. Chairman, I appreciate the opportunity to submit a statement for the record today to provide a business perspective on the recommendations for establishing a voluntary, national system of educational standards and assessments.

The National Alliance of Business reiterates its support for including in H.R. 4323 the recommendations contained in the National Council on Education Standards and Testing's final report entitled *"Raising Standards for American Education."* Our support for the recommendations recognizes that there will be many practical issues to work out during a process of development and implementation. But, business support for action to begin the process is broad and strongly held.

As you know, the Business Coalition for Education Reform, composed of 11 key national business organizations, urged the House committee in a January 27 letter to incorporate the recommendations of the Council in their education reform bill, H.R. 4323. In addition, last week, both the National Alliance of Business Board of Directors and the advisory council to the Alliance's Center for Excellence in Education were unanimous in their support for the Council's recommendations and in their call for including such provisions in the House and Senate bills. (The advisory board of the

BEST COPY AVAILABLE

Alliance's Center for Excellence in Education includes many of the nation's educational leaders, including the presidents of the two largest teacher unions -- list attached.)

We fully recognize the complexity of the issues and the unresolved questions that raise concern for some Committee Members, but we are convinced an effective national entity can be constructed by the Congress, and a process started, in which these problems can be discussed and resolved. We should not lose this opportunity to take our best shot at developing educational standards and assessments in a basic education reform bill that may last a decade. It will take time and experimentation, but we must begin the process. The recommendations provide a formal, long-term process to assist in reviewing systemic education improvements. By including the Council's recommendations in H.R. 4323, the Congress can add its authority and intent to this vital activity.

I encourage the Committee to add whatever appropriate safeguards it feels is necessary to ensure that the development of standards and assessments is equitable and practical. No one can anticipate and resolve all of the problems before we proceed to act. Solving the problems that can be identified along the way is part of what the process of developing standards and assessments should address. The critical point I wish to emphasize is that we should not lose this opportunity to proceed on the recommendations while we focus on the details of potential problems that cannot be quickly resolved. It is far better, in my view, to get started with the full knowledge that we will have to fix problems along the way than not to start at all.

The endorsement of the Council's recommendations is based on our belief that they provide an indispensable starting point, and a foundation, for developing a voluntary system of standards and assessments necessary to improve the quality of American

education. The business community believes that a voluntary system of educational standards and assessments is a critical component not only of education reform but also of our competitive and economic vitality as a nation. All of our key competitor nations have found ways of establishing such systems and of applying them so that virtually all citizens benefit. The point our business colleagues make is that we must act now, and at least begin the long process of developing fair and effective standards and assessments.

One of the commonly misunderstood areas of policy and practice that would benefit from national leadership is the area of assessments. The recommendations of the Council are compatible with the National Alliance of Business position statement on assessment -- developed last July with the help of the Alliance's Center for Excellence in Education advisory Board members -- which states in part:

> *The National Alliance of Business supports the establishment of a national assessment system for America's elementary and secondary students. This system would consist of 1) definitions of what students should know and be able to do, 2) standards of desired performance set at internationally competitive levels, and 3) varied and multi-year assessments — such as portfolios, exhibitions, performance exams, and multiple choice tests — focused on mastery of broad concepts and bodies of knowledge, higher order thinking skills, and interdisciplinary problem solving.*

and further,

> *The Alliance recognizes that developing a national assessment system is a complex endeavor. Consideration must be given to issues of equity, cost requirements, the relationship of these assessments to existing standardized*

tests, the reporting and actual use of the results, and the intervention strategies needed to assist those schools and students who at first will not do well on the assessments. Three major developments must take place before a national assessment system can be put in place: 1) greater clarity regarding what we want to assess, 2) investments in curricula, pedagogy, and staff development designed to enable students to learn such things, and 3) the development of wholly new kinds of assessment instruments sensitive to these new topics and competencies."

(The complete text of the National Alliance of Business statement on assessment is attached.)

The Alliance believes that a unified set of voluntary educational standards can assist state and local educators in their efforts to restructure educational systems and to provide better learning opportunities for all youth. We think it is important for the Congress to be involved in monitoring progress toward the national education goals, in setting and certifying voluntary education standards, and in building credible assessment systems for education. We recommend that H.R. 4323 should:

- Make permanent the National Education Goals Panel and expand its membership to include members of Congress.

- Establish a permanent national entity to:
  - -- coordinate the development of voluntary national education standards for both students and school delivery;
  - -- coordinate the development of a system of national assessments for student performance;

-- ensure that assessment systems are valid and that they measure the established and agreed upon standards; and
-- provide a forum for expressing diverse opinions and knowledge about standards and assessments so that critical issues can be examined and consensus built to resolve them.

From our perspective, the act of implementing the Council's recommendations will provide some of the tools necessary to make systemic educational change. And these recommendations cannot be implemented effectively without the support and involvement of Congress.

The National Alliance of Business is committed to working with the Congress to incorporate these recommendations in final legislation.

Attachment

POLICY STATEMENT OF THE NATIONAL ALLIANCE OF BUSINESS ON A NATIONAL ASSESSMENT SYSTEM

July 1991

Preamble

Attachment

The National Alliance of Business is convinced of the importance of a *national assessment system*. By that we mean *more than a single test* at the national level that tells all of us how well our students are progressing against well-defined national standards. National benchmarks are important, but what is really needed is an *entire system that provides coherent guidance, information, and accountability at all levels on how well our students, schools, districts, and states are doing in bringing students to the levels of achievement critical for successful citizenship and employment.*

Some argue that a national system would undermine the benefits derived from our federated system of education. We agree that this is a diverse country. However, we feel certain that there are certain skills and abilities that *all* students need if they are to be successful in society. A system that captures how well all our children are gaining those skills and abilities is not only appropriate, but important. At the same time, states, districts, and schools can continue to capture additional information they hold important.

Ideally, the assessments developed for teachers and schools to measure individual student achievement -- in all its complexities -- should be rolled up to determine national progress, compare progress among the states, and compare our nation to others. To be sure, there are many concerns about whether this is feasible. But it is important enough that our nation, with the active investment of the federal government, should try to make such a system a reality. Faulty assessments can do more damage than good to the quality of education.

Position Statement

The National Alliance of Business supports the establishment of a national assessment system for America's elementary and secondary students. This system would consist of 1) definitions of what students should know and be able to do, 2) standards of desired performance set at internationally-competitive levels, and 3) varied and multi-year assessments -- such as portfolios, exhibitions, performance exams, and multiple choice tests -- focused on mastery of broad concepts and bodies of knowledge, higher order thinking skills, and interdisciplinary problem solving.

The Alliance only supports such an assessment system if it is used to bring *all* students to a high level of achievement, rather than relegating some students to lesser quality educational programs or permanently excluding them from higher education or economic opportunities. More resources, not less, should be devoted to those students who initially have difficulty meeting the international standards. Care should be taken to ensure that the assessment is not biased in ways that discriminate on the basis of race or ethnicity.

The development of the national assessment system should be coordinated by an *independent, national* entity, or commission, established in federal statute and supported by federal funds. This entity would be responsible for convening representatives of interested and affected groups -- including academicians, politicians, educators and teachers, business people, community organizations, and parents -- in a collaborative process to develop the definitions and achievement standards. It would also oversee the development and implementation of the assessments, and the eventual reporting of assessment results.

The national definitions and standards should inform parents and the public of the knowledge and skills students would be expected to master and the standards to which they would be held. They must not, however, constitute a national curriculum with specified textbooks and instructional strategies. States, districts, schools, and teachers must have the flexibility to choose their own textbooks and design their own instructional strategies for achieving the national standards and accommodating racial, ethnic, and geographic diversity.

States and districts (or state and district consortia) could develop their own assessment instruments for measuring student proficiency. The national commission would oversee and support these development efforts, and provide for their calibration to common standards for comparative purposes nationally. The national commission would also coordinate the development efforts, facilitating the sharing of information and preventing unnecessary duplication of effort (and waste of money).

National assessment should serve two purposes: 1) as a benchmark for measuring national progress against defined standards over a period of years, and 2) as a means for guiding curriculum and instruction so that *all* students obtain the skills and knowledge they need to participate fully as citizens, workers, and consumers. Because of this dual purpose, the Alliance recommends that a two-pronged approach be used to develop the assessment system, with both "prongs" being based on the same definitions and standards.

First, existing assessments that rely heavily (though not exclusively) on multiple choice tests (such as NAEP, the National Assessment of Educational Progress) should be used to assess students' skills against as many of the standards defined through the collaborative process as possible. These assessments should be adapted and expanded to measure as many of these standards as possible, although they cannot be expected to measure *all* of the skills students will have to obtain, such as

the ability to speak clearly and effectively, to work productively in groups, or to solve "ill-structured" real world problems. However, they could serve as benchmarks of national, state, and possibly district progress against many of the defined standards. Additionally, they could be developed within a relatively short period of time, be delivered on a sample basis and at a relatively low cost, and be used in international comparisons.

Second, a more comprehensive assessment system that accurately assesses *all* of the skills and knowledge we want students to possess also should be developed. This assessment could not be a single paper-and-pencil test. It would have to employ more complex techniques such as portfolios, exhibitions, and performance exams, along with some multiple choice tests. Initially, its primary use would be to inform and guide teachers in the preparation of their students. Over time, the national commission should develop a method of "rolling up" students' scores on these assessments so they could replace the less comprehensive assessment described above as a means for measuring national progress.

Because it would assess *all* of the skills and knowledge we want students to possess, this comprehensive assessment *should* be used to guide curriculum and instructional practices in the country. This can be done by making the assessment "high stakes" -- making performance on it affect important outcomes. Colleges and employers should use students' performance results on this assessment as one of the criteria in the selection of entrants to their institutions, thus making it serve as an incentive to both students and schools.

Furthermore, this more comprehensive assessment system should be used as an accountability mechanism for the education system. States, districts, schools, and even school professionals could be held accountable for improvements (not absolute levels) in their students' performance on the assessments. The assessments could be used to identify schools and school professionals who help students to succeed against the standards, and inform district administrators and the public about schools needing help to improve. It should provide a fair means for taking district- or state-level action where failure to move toward the standards continues.

The Alliance recognizes that a national assessment system alone cannot bring about the high educational achievement of students. For the assessment system to be effective, it must be part of a comprehensive and systemic approach to restructuring education. Teachers and their subject matter associations (e.g. math and science associations) should take the lead with researchers, business people, other educators, and community leaders in national efforts to develop new curricula, prepare materials, and redesign teaching strategies. Colleges and universities must revamp teacher preparation programs, and they, state and local education administrators, and teachers must develop new and extensive staff development and training programs for existing teachers that prepare them to work within the new system. Teacher preparation and training are often neglected when education change initiatives are launched; the Alliance wants to make clear that such programs must

receive priority attention if the new assessment system is to be successful in improving student achievement. Furthermore, if schools and school professionals are to be held accountable for educating students to the standards, a system of site-based decision making must be part of the process.

The Alliance recognizes that developing a national assessment system is a complex endeavor. Consideration must be given to issues of equity, cost requirements, the relationship of these assessments to existing standardized tests, the reporting and actual use of the results, and the intervention strategies needed to assist those schools and students who at first will not do well on the assessments. Three major developments must take place before a national assessment system can be put in place: 1) greater clarity regarding what we want to assess, 2) investments in curricula, pedagogy, and staff development designed to enable students to learn such things, and 3) the development of wholly new kinds of assessment instruments sensitive to these new topics and competencies.

Appropriate and adequate assessment procedures do not exist at this time, and it will require a multi-year effort to develop them for the various subjects, skill areas, and grade levels. However, for a few of these areas a great deal of preliminary work has been accomplished already, and we should proceed immediately and aggressively to develop the desired assessments in those subjects within the next two to three years.

Development of a national assessment system must be seen as a national priority by the Congress and the Administration. It should be carried out with all deliberate speed, but in such a manner as to insure that it works effectively the first time around. The American business community should actively support and encourage this process.

CENTER FOR EXCELLENCE IN EDUCATION

The NAB Center for Excellence in Education's Advisory Board is composed of nationally recognized education and business leaders who provide the Alliance with practical advice on education policies and programs.

CO-CHAIRS:
Edwin D. Campbell
Executive Director
Coalition of Essential Schools
Brown University

John E. Pepper
President
The Procter & Gamble Co.

MEMBERS:
Gordon M. Ambach
Executive Director
Council of Chief State School Officers

Sue E. Berryman
Director
Institute on Education and the Economy
Teachers College
Columbia University

Ernest L. Boyer
President
The Carnegie Foundation for the Advancement of Teaching

James B. Campbell
Chairman of the Board
MISSCO Corporation

James A. Cannavino
IBM Vice President and General Manager, Personal Systems
IBM Corporation

Michael N. Castle
Governor
State of Delaware

Malcolm B. Chancey, Jr.
President
Liberty National Bank and Trust Company of Louisville

Constance E. Clayton
Superintendent of Schools
School District of Philadelphia

Denis Doyle
Senior Research Fellow
The Hudson Institute

Marian Wright Edelman
President
Children's Defense Fund

Joseph A. Fernandez
Chancellor
New York City Schools

Calvin M. Frazier
Professor and Director
Center for Educational Policy Studies
University of Denver

Martha C. Fricke
Past President
National School Boards Association

Keith Geiger
President
National Education Association

Jack L. Heckel
President and COO
GenCorp

Thomas R. Hefty
President and CEO
Blue Cross & Blue Shield United of Wisconsin

Bill Honig
California Superintendent of Public Instruction

David W. Hornbeck

Malcolm R. Lovell, Jr.
Distinguished Visiting Professor and Director, Labor Management Institute
George Washington University

William L. Lurie
President
The Business Roundtable

Pedro F. Mata
President
Grace Cocoa

Harold W. McGraw, Jr.
Chairman Emeritus
McGraw-Hill, Inc.

Deborah Meier
Principal
Central Park East Secondary School

Frank Newman
President
Education Commission of the States

Kenwood C. Nichols
Vice Chairman
Champion International Corporation

James F. Orr, III
Chairman and CEO
UNUM Corporation

Matthew Prophet
Superintendent
Portland Public Schools

Ian M. Rolland
President and CEO
Lincoln National Corporation

Raymond C. Scheppach
Executive Director
National Governors' Association

Albert Shanker
President
American Federation of Teachers

Franklin L. Smith
Superintendent
District of Columbia Public School System

Marc Tucker
President
National Center on Education and the Economy

Ronald A. Wolk
President
Editorial Projects in Education
Education Week

William S. Woodside
Chairman
Sky Chefs, Inc.

Alan L. Wurtzel
Chairman
Circuit City Stores

Statement of

Eva L. Baker
Professor, University of California, Los Angeles

Robert L. Linn
Professor, University of Colorado

Co-Directors
National Center for Research on Evaluation, Standards, and Student Testing
(CRESST)

before the Subcommittee on Select Education
Committee on Education and Labor
U.S. House of Representatives

March 17, 1992

Mr. Chairman, thank you for the opportunity to testify on the topic of needed research in educational assessment. My name is Robert Linn. I am a professor at the University of Colorado and co-director with Eva Baker of the National Center for Research on Evaluation, Standards, and Student Testing (CRESST). Although she was scheduled to speak today, I was able to save resources by substituting for her. This testimony is provided on behalf of us both.

We wish to address our remarks to the need for expanded research in the area of assessment. It is not surprising to hear of the importance of research from a researcher. But we draw our position from a larger context—from the public discussion surrounding educational reform. In the deliberation of the National Council for Standards and Testing, in the debate among participants on the Council's Task Force on Assessment, and in the summary of the report released by the Office of Technology Assessment on Testing, it is clear that the public, educators, and policymakers have multiple expectations for tests and

make many assumptions about the validity of the information tests and assessments provide.

These interested parties want assessment for accountability, for instructional improvement, for certifying student accomplishments, and they want them now.

A case in point is the national attention to assessments of student performance, using projects, experiments, essays, and portfolios instead of more familiar types of testing. CRESST researchers will provide information on the validity of some of these tests for different uses, and guidelines to develop and validate these measures. We already have results in CRESST's research on history and science assessments that we believe will make performance assessment significantly more cost effective and valid. But in order to begin to approximate the appetite of the public for test information--and to assure for ourselves and others that testing supports rather than impedes student learning, additional questions require sustained attention.

Let us list a few.

How fair are performance assessments to children of different backgrounds? How does gender impact performance when new types of assessments are used? How do we assure that economically disadvantaged or limited English-Proficient, or special populations, such as the learning disabled, are fairly treated?

How is fairness influenced by different administrative procedures and scoring procedures? How should comparability of design or administration be judged? How should comparability of test results be determined?

What is the impact of various programs of instruction on these new tests? Do the assessments measure the things we wish to teach? What kinds of important learnings can be measured? How general or transferable is performance from project to project or test to test or to other important

accomplishments? How do these assessments predict readiness for the workforce or for postsecondary education?

Are performance tests less or more corruptible than traditional tests? How trustworthy are our findings and what steps can increase our confidence in test performance?

Can multiple purposes of assessment be simultaneously served with validity? Can an assessment contribute to teaching and learning and accountability? What is the impact of new assessments on students and on the quality of life in school?

What are the best ways for teachers to be involved in the design and use of new assessments? What ways are cost effective? What ways contribute directly to student learning? How can the design of new assessments build upon our knowledge of thinking and learning?

How should we measure students' ability to integrate across subject matter areas? How should we measure their effort, thinking processes, and habits of mind?

How should we integrate the range of assessment options before us to provide the best information, with the most positive consequences and within reasonable costs?

Beyond the solutions to these problems, we must develop improved ways to communicate the results of student assessments. Many believe that testing has developed in the manner it has in part because of the public's desires for easy, understandable answers—answers that depend on numbers. Changing the basis of assessment to more concrete analyses of children's performance may very well change the kind of information policymakers and the public receive. Research on how to communicate complex information is essential if real improvements in assessment are to be possible. We must show details and realities of what children can do rather than exclusively rely on abstractions, like scores and averages. We will also find ways to reach out to all parents who need to have good information about their schools—and who need that information in a form that they can use to support children's growth and accomplishments.

Most of these questions are under study now. Some of them will take considerable time to answer. The questions listed share a focus on student outcomes. However, we believe we must make a substantial effort to address another class of questions about educational reform. Especially if educational reform is to be systemic, we must focus the attention of scholars and talented practitioners on how to assess the quality of schooling itself. Clearly, under certain conditions, we can infer something about instruction from student performance. But many of us worry that we have not paid enough attention to the description of student experiences and school environments. Why should we expect children to do well in school—and have high test scores—if their schools may not be safe? Why should we expect children to excel when they may not have challenging textbooks, or enough of any books to go around? How can they have homework without paper? In order to understand the results from any tests, whether multiple-choice, or performance-based, we need to be able to make accurate statements about what school experiences are like and how they relate to student outcomes. If we cannot, we will never solve the equity problem.

During the deliberations of the National Council on Education Standards and Testing, issues of school experiences and environments were discussed in at least three different ways. First, they were discussed in the light of "delivery standards" for the schools and for systems. Secondly, they were discussed as part of needed evidence before assessments were to be certified for use for a "high stakes" purpose, such as to promote or graduate a student. Third, delivery standards were discussed as necessary to assure the implementation of educational reform. Delivery standards were an extremely controversial topic. To some, they implied a prescriptive or controlling function, a way to homogenize

schools and classrooms from inside the beltway. To others, they conjured up horrors of more checklists, paperwork, and mandated but unread reports.

We must be able to conduct research on the assessment of delivery standards—on school experiences—to determine if we can develop good, cost-effective information that helps schools reach their potential and serve the interests of their students. If we do not undertake this research, we will perhaps always have nagging questions about the quality of our judgments and the fairness of decisions made on student outcomes alone.

The National Council on Education Standards and Testing also focused on the importance of assessing the impact of the spate of assessment activity. The idea, a long-term, independent study of new testing policies should be strengthened and preserved in the OERI.

As the Office of Technology Assessment Summary of the Report, *Testing in American Schools: Asking the Right Questions* observes, "Congress has an important role to play in supporting R&D in educational testing, because adequate funding cannot be expected from other sources" (1992, p.36). Because the testing and assessment agenda is essential for understanding educational quality, we must be sure that long-term, Federally-supported programs of research on assessment are continued.

National R&D Centers

All research must be conducted by a mix of research providers: individual researchers, commercial companies, and state agencies. But we wish to speak

about the importance of preserving the programmatic research orientation of National R&D Centers—Centers which have made and can continue to make important contributions to knowledge and practice in assessment, as well as in areas such as school organizations, learning, policy, and teaching. What a Center on assessment does is to serve as a trusted source for impartial analysis on important issues, and as a focal point for creative research to help solve and understand our problems. CRESST regularly fields requests from Congressional staff, members of the Administration, State legislators, the press, educational institutions, and individual parents, teachers and representatives of the business community. They want to know the state of knowledge in assessment. On last Friday, one of us had calls from the press, the Arts community on alternative assessment, a technology company on portfolios, two non-profits, asking about workforce readiness, a school psychologist, a teacher, and a graduate student about what we knew and where they could go for additional help. The other of us was meeting with the Math Science Education Board's Study Group on Guidelines for Mathematic Assessment. During the week, we talked to school district and state administrators, university researchers at many sites, planned a principal's workshop, an international meeting at UCLA on assessment, a seminar for state legislators, met with the Chicago educational reform community, and finished four progress reports due to OERI. Saturday, we presented our results on what we have learned about performance assessment to more than 200 teachers. Today, one of us will meet with colleagues from the State Alternative Assessment Exchange and the National Assessment Governing Board to work on problems of linking and comparing results from different assessment systems. And that was just one week. CRESST is called not because its single purpose is dissemination, it is called upon because the quality of its research productivity is valued.

Individual researchers alone, working on small projects, simply cannot meet needs of this sort. The bombardment of questions from the field keeps us close to real problems and issues on assessment, though housed in universities we may be, and allows us to adjust research programs in the light of our data as well as in terms of the practical problems we confront daily. Furthermore, because we are researchers, we are careful to give information within the boundaries of our knowledge.

The idea of a Center program is important and it works; it should be retained within the newly proposed OERI structure for assessment, as well. CRESST's strength resides in the ability of teams of researchers to propose their own theories, strategies, and methods within the guidelines of expectations from the OERI, informed by what they learn from research and by their interaction with the field. The capacity and sustained attention to complex assessment research problems requires long-term programs and significant resources.

In summary, we believe that educational assessment should be a significant part of the research plans of OERI. We believe that we all must address issues of the assessment of individuals and schools—measuring standards for students and institutions—to be sure that our expectations for assessment work for the good of children and for the future of our country. We also believe that for all the range of issues identified by the proposed OERI institutes, and particularly for assessment, long-term, programmatic research by university-based R&D Centers should be an essential element in the American R&D network.

National Association of State Boards of Education
1012 Cameron Street
Alexandria, VA 22314
(703) 684-4000
FAX (703) 836-2313

President
Past President
President Elect
Vice-President
Secretary-Treasurer
Northeastern Area Directors
Southern Area Directors
Central Area Directors
Western Area Directors
Pacific Islands Area Director

April 7, 1992

Representative Dale E. Kildee
United States House of Representatives
2239 Rayburn House Office Bldg.
Washington, D.C. 20515-2207

Dear Representative Kildee,

On behalf of state boards of education, I would like to take this opportunity to respond to the move by Congress to adopt the recommendations made by the National Council on Education Standards and Testing, (NCEST). NASBE believes that provisions for the development of voluntary, national standards and testing should be included in H.R. 4323. While we offer our support, we do so with concern for we realize we are making decisions that will affect the lives of our most precious resources, our youth. We feel that there are a number of critical issues that must be addressed in this discussion.

We offer the following position, formulated and adopted by NASBE's Governmental Affairs Committee, an 18 member politically and geographically diverse body. The position was also approved by the NASBE Board of Directors, also made up of state board of education members.

It is clear that standard setting and testing in American education is going to be reshaped by the Congress. However, there are a number of questions and issues that need to be resolved or responded to around the NCEST report and the role of the National Assessment of Educational Progress in the legislation.

NASBE POSITION ON STANDARD SETTING AND ASSESSMENT

Student Performance Standards

1. We support a national effort to create performance standards in all key subject areas including both standards for content and student capability.

2. Planning and decision-making for the system should be characterized by:

   - o Substantial involvement of leading experts and organizations in each area of curriculum;
   - o Substantial involvement of teachers and school administrators (teachers should be involved in designing and using the assessment system);
   - o Substantial opportunity for input from parents, the general public, and state policymakers;

Gene Wilhoit
Executive Director

- o Careful attention to drawing on the best existing development efforts by states, local school districts and discipline-based organizations on student outcomes; and

- o Sufficient timelines to allow several cycles of broad feedback on draft standards before final decisions are reached.

3. Based on the student performance standards a parallel system of indicators will be developed on "opportunity to learn" or school delivery standards to assure that eventual assessments of schools takes into account the students opportunity to receive instruction by well trained teachers, with quality materials and technology. We support the development of school delivery standards by the states because we feel that assessment results should be reported in the context of other relevant information. A number of primary factors such as context (type of community, socioeconomic status of students, and school climate), and resource (expenditures per students, staffing) have been discussed by some as key indicators. However, we feel it is equally appropriate to focus on program and process factors that include areas such as quality of curriculum, instructional methods and outcomes (student performance, dropout rates).

The issue of who develops school delivery standards is by far one of the most politically charged issues eminating from this discussion and there is no simple resolution. We feel that such standards should be included in this process. We feel that Congress must reach some compromise on the issue that recognizes state autonomy but that also guarantees comparability.

4. The governance and implementation of performance standards must include state policymakers.

National Assessment of Students

Our overriding foundation for the support of any assessment system is that the primary purpose of the assessment system should be to help educators and policymakers through improved instruction and by advancing student learning. To this end the following steps must be taken:

1. Substantial long-term investment in research and development is needed on the costs, technical validityand reliability, lack of cultural, racial and sexual bias and quality administrative requirements for systems of tests designed to measure student performance on national content standards.

2. Resources should be provided from the federal government to states and consortia of states to develop systems of tests around content standards, including extensive consultation with experts, educators and the public; careful field-testing and revision of tests; and widespread training of teachers in test administration.

3. No federal funding should be tied to the outcomes of these tests and there should be no public reporting of results or use of tests for high stakes decision-making until the accuracy and credibility of the assessments is established to expert and public satisfaction.

4. There should be no public reporting of results without simultaneous data on the extent to which students have enjoyed equal opportunity to learn content measured by the assessments.

5. There must be federal support for costs of local tests administration and analysis and reporting.

Page 3

Questions and Issues for Discussion

1. The foremost critical issue raised in the development of a national test is that of costs. Will the development of national assessments and measures be worth the effort and expense? In comparison to other obvious needs to reform education (equitable distribution of money, health care, Head Start, or teacher development) should we be expending scarce funds for national assessment? Can we afford these new sorts of tests and who will pay for them? Since these newly designed schemes will cost approximately 10 times our current norm reference standardized test, are they doable? We urge that these efforts not come at the expense of other crucial federal programs.

2. A number of issues must be addressed about the use of the test. Is the intent to monitor school-performance or will it be used for high stakes purposes such as graduation, college entrance and employment? How will Congress assure the appropriate use of these materials in this highly political environment?

3. The current assumption is that the focus should be on subject matter competency in english, math, science, history, and geography. Is this satisfactory? We have some concerns that the focus on these national standards might narrow the work of State Boards of Education that reflect a broader base of subjects, higher order thinking, integrated learning, work force skill development, virtues, character and citizenship.

3. There are concerns about assessment as the appropriate lever for reform. Will this assessment system boost academic achievement? How effective are assessments in improving teaching and learning? Will this national assessment lead to a national curriculum? Will this suppress intellectual freedom? Can top-down assessment affect true reform at the school level?

It would be irresponsible for Congress to commit itself too hastily to these concepts without setting into place a means for questioning and resolving them.

4. Finally, a number of concerns have been raised about fairness. Will this national assessment and testing program negatively impact the richness of diversity in this country? How will this national assessment program affect people of color and the disadvantaged?

On the related issue of NAEP the Office of Technology Assessment, (OTA) concluded that Congress should weigh very carefully, any proposals to change NAEP and focus on ways to retain and strengthen it as a national indicator of educational progress. We share OTA's concerns and urge that NAEP not be administered to every child or be used as a basis for educational decisions about children and schools.

Congress may decide to take advantage of this opportunity to focus on a long-term education reform effort, to establish a national entity which can provide a forum and a process for developing such standards and assessments. But, we urge that Congress ensure that the system developed be fair and effective.

Thank you for this opportunity to comment.

Sincerely,

Gene Wilhoit
Associate Director

ISBN 0-16-039802-9